Proceeding of the 1st International Conference of the Moroccan Association of Operational Research and Decision Support, ROADEM’24

Preface

Welcome to the *Proceedings of ROADEM'24 Conference*. This volume compiles the latest research, innovations, and discussions from the 2024 edition of the ROADEM Conference, a premier gathering for experts, scholars, and professionals in the fields of Operational Research and Decision Support.

ROADEM'24 has once again served as a platform for presenting cutting-edge ideas and fostering interdisciplinary dialogue among attendees from around the globe. We are proud to feature a diverse array of papers that explore a wide range of topics, from logistics and transportation. Each contribution represents a significant step forward in understanding and addressing the challenges facing our industries today.

The success of the conference would not have been possible without the dedication of our esteemed authors, reviewers, and organizers. We extend our heartfelt gratitude to everyone who has contributed to making this event a success. Their hard work and passion for advancing knowledge have made this compilation of research possible.

We hope that the discussions and insights presented in these proceedings inspire continued progress and innovation within the field. As the world evolves, so too must our solutions to the challenges ahead, and it is through collaboration and knowledge-sharing that we can create sustainable, effective, and impactful solutions.

We thank all attendees, speakers, and contributors for their active participation and look forward to continued growth in the years to come.

Enjoy the proceedings, and we wish you all a fruitful journey of discovery and inspiration.

Dr. Fatima Bouyahia
Cadi Ayyad University, Morocco

Prof. Abdelhamid Benaini
Le Havre University Normandy, France

Dr. Jaouad Boukachour
Le Havre University Normandy, France

A Hybrid Approach to Solve Quay Crane Scheduling Problem in Container Terminal

AIDI Sanaa[1], Mazouzi Mohamed[2]

[1] Affil1University Hassan 2, national higher school of electricity and mechanics, Casablanca,
Department of Mechanics, engineering and innovation, Morocco
Sanaa.aidi@gmail.com

2 University Hassan 2, national higher school of electricity and mechanics, Casablanca,
Department of Mechanics, engineering and innovation, Morocco
[2] Affiliation
m.mazouzi@ensem.ac.ma

Abstract

The Quay Crane Scheduling Problem (QCSP) is a critical challenge in optimizing operations within container terminals, significantly impacting vessel turnaround times and overall efficiency. This study proposes a hybrid algorithm that combines genetic algorithms with local search techniques to address the complexities of the QCSP. The methodology begins with the generation of an initial population of feasible schedules, followed by a selection process based on makespan minimization. Crossover and mutation operations are employed to create new candidate solutions, which are then refined through local search strategies. Experimental results demonstrate that the hybrid approach outperforms traditional methods, such as pure genetic algorithms and tabu search, in terms of makespan reduction and resource utilization. The findings indicate that this hybrid method not only enhances operational efficiency but also provides a robust framework for adapting to dynamic terminal environments. This research contributes to the field of logistics and operations management by offering a practical solution to improve scheduling efficiency in container terminals.

Keywords: *Quay Crane Scheduling Problem, Hybrid Algorithm, Genetic Algorithms, Local Search, Container Terminals.*

1 Introduction

The Quay Crane Scheduling Problem (QCSP) is a critical aspect of container terminal operations, significantly impacting the efficiency of cargo handling and overall port productivity. As container terminals serve as vital nodes in global supply chains, optimizing the scheduling of quay cranes is essential to minimize vessel turnaround times and enhance operational efficiency [1]. This article explores a hybrid approach that combines mathematical optimization and simulation techniques to effectively address the QCSP[2]. The QCSP involves determining the optimal sequence of operations for a set of quay cranes tasked with loading and unloading containers from vessels. The primary objectives include minimizing the makespan (total time required to complete all tasks), reducing delays, and ensuring the effective utilization of crane resources. The complexity arises from constraints such as non-interference between cranes, safety distances, and task priorities based on cargo urgency. Recent research has proposed a Hybrid Genetic Algorithm (HGA) as an effective solution for the QCSP. This method leverages the exploratory capabilities of genetic algorithms while incorporating refined local search strategies to enhance convergence speed and solution quality. The HGA begins by generating an initial population of feasible schedules, ensuring compliance with operational constraints through heuristic approaches. The selection process favors solutions with shorter makespans, iteratively refining offspring solutions to converge towards optimal schedules[3].

In addition to HGA, simulation techniques play a crucial role in managing uncertainties and dynamic changes within container terminal operations [4]. By simulating various scenarios, terminal managers can evaluate the impact of disruptions such as equipment failures or sudden changes in cargo handling requirement on crane schedules [5]. This allows for timely adjustments and rescheduling strategies that maintain operational efficiency under varying conditions.

Extensive computational experiments using established benchmark datasets have demonstrated the effectiveness of hybrid approaches in solving the QCSP[6]. For instance, studies have shown that employing HGA can lead to significant reductions in makespan and improved resource utilization compared to traditional scheduling methods. Moreover, integrating simulation techniques allows for better preparedness against unforeseen disruptions, thereby enhancing overall resilience in terminal operations.

The hybrid approach combining genetic algorithms with simulation techniques offers a robust framework for addressing the complexities of the Quay Crane Scheduling Problem in container terminals [7]. By optimizing crane schedules while accounting for potential disruptions, this methodology not only enhances operational efficiency but also contributes to the seamless functioning of global supply chains. Future research may focus on further refining these algorithms and exploring their applicability across different terminal environments.

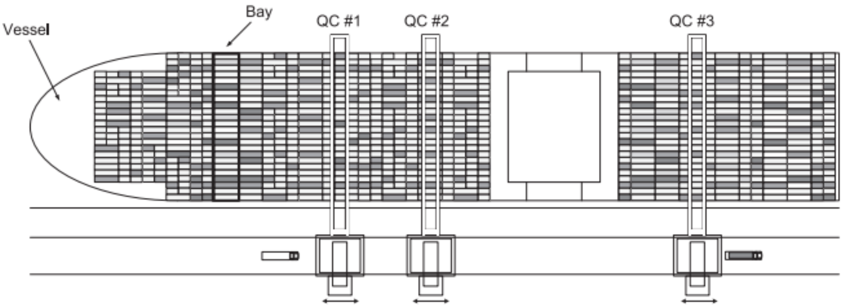


FIG. 1:Illustration of QCs working on a vessel

2 Literature Review

The Quay Crane Scheduling Problem (QCSP) is a critical area of research in operations management, particularly within the context of container terminals[8]. As container traffic continues to grow globally, optimizing the scheduling of quay cranes becomes essential for enhancing terminal efficiency and reducing vessel turnaround times. This literature review synthesizes key findings from recent studies [9], highlighting various approaches and methodologies employed to tackle the QCSP.

QCSP involves determining the optimal sequence and timing of operations for quay cranes tasked with loading and unloading containers from vessels. The problem is characterized by its complexity due to constraints such as crane interference, time windows, and dynamic operational environments influenced by disruptions [10]. The goal is to minimize the overall makespan while maximizing resource utilization.

2.1 Methodological Approaches

Several methodologies have been proposed to solve the QCSP, ranging from exact algorithms to heuristic and metaheuristic approaches:

2.1.1 Exact Algorithms:

- Mixed Integer Linear Programming (MILP) and dynamic programming have been employed to derive optimal solutions for small to medium-sized instances of the QCSP. For example, a study at the Port of Tripoli utilized these methods to minimize loading and unloading times effectively[11].
- Branch and Bound techniques have also been applied, demonstrating effectiveness in solving static instances of the problem but often struggling with dynamic conditions [12].

2.1.2 Heuristic Methods:

- Genetic algorithms have gained popularity due to their ability to provide near-optimal solutions within reasonable computational times. These methods are particularly useful in scenarios where exact solutions are computationally infeasible [6].
- Tabu search heuristics have been developed to address specific features of the QCSP, such as time windows and crane interference, showing promising results in improving solution quality [9].

2.1.3 Hybrid Approaches:

- Recent studies advocate for hybrid methodologies that combine various optimization techniques. For instance, integrating genetic algorithms with local search strategies has shown significant improvements in solution quality for dynamic scheduling problems[1].
- A notable example is the Hybrid Genetic Algorithm (HGA), which combines genetic algorithms with simulation techniques to manage uncertainties effectively in quay crane scheduling.

The literature on the Quay Crane Scheduling Problem reflects a rich tapestry of methodologies aimed at optimizing terminal operations. While significant progress has been made through exact algorithms, heuristics, and hybrid approaches [13], ongoing research is essential to address the complexities presented by dynamic environments and real-world applications [14]. Continued exploration into adaptive rescheduling strategies will further enhance the efficiency of container terminals worldwide.

3 Methods

Mathematical modelling of the Quay Crane Planning Problem (QCSP) is essential for optimizing operations in container terminals. This complex problem involves the coordination of several cranes to load and unload containers while respecting various operational constraints. This section presents an overview of the mathematical modelling approaches used in the literature to tackle this problem.

3.1. Mathematical formulation

Decision variables: These variables represent decisions to be made, such as the assignment of cranes to vessels and the sequence of operations. For example, a binary variable Xij could indicate whether crane i is assigned to ship j.

- Objective function: The main objective is usually to minimise the total operating time or makespan. This can be expressed as follows :

$$\text{Minimize } Z=\sum Cj \quad (1)$$

where Cj is the completion time for each ship j.

- Constraints : Constraints are essential to ensure the feasibility of solutions. They may include

- Non-interference: Cranes must not interfere with each other during simultaneous operations.

$$xij+xik\leq 1 \forall i\in G,\forall j,k\in V,j\neq k \text{ et } tij+dij>tik \text{ ou } tik+dik>tij \quad (2)$$

- Capacity: Each crane has a maximum capacity of containers it can handle at any one time.

$$\sum cijxij\leq Ci \qquad \forall i\in G \quad (3)$$

- Waiting times: Waiting times for ships must be minimised.

$$\sum(tj^{wait})\leq T\max \quad (4)$$

$$tj^{wait}=tj^{start}-tj^{arrival} \quad (5)$$

3.2 method of resolution

Several methods have been developed to solve the QCSP:

1. Linear programming: MILP formulations are solved using optimization solvers such as CPLEX or Gurobi. These tools can find optimal solutions for moderately large instances.

2. Heuristics and metaheuristics: Given the NP-hard complexity of QCSP, heuristic approaches such as genetic algorithms, tabu search, and ant colony algorithms are commonly used. These methods provide approximate solutions in a reasonable time .

3. Hybrid approaches: The combination of optimization and simulation is an emerging trend in QCSP research. For example, the use of a rolling horizon framework allows plans to be dynamically adapted to operational uncertainties.

Proposed hybrid method

The proposed hybrid method is based on integrating a genetic algorithm with an improved local search. The key steps are

1. Initialization

- Generation of an initial population of crane schedules using heuristics to ensure that all solutions respect operational constraints.

2. Selection:

- Solutions are selected on the basis of their quality, measured by their makespan. The best solutions are kept for reproduction.

3. Crossover and mutation:

- Application of crossover and mutation operations to create new solutions from the selected best solutions.

4. Local search:
   - A local search is applied to the new solutions to further refine the schedules, ensuring that the algorithm converges to optimal or near optimal solutions without excessive computational cost.
5. Evaluation and Iteration:
   - The new solutions are evaluated and the process is iterated until a stopping criterion is reached, such as a maximum number of iterations or a minimum improvement in the average processing time (makespan).

4.Results and validation

Computational experiments were carried out using established benchmark datasets. The results show that the Hybrid Genetic Algorithm (HGA) significantly reduces the makespan and improves the resource utilization of the cranes compared to traditional methods (Aidi et al., 2025).

The hybrid approach, which combines a genetic algorithm with a local search, provides a robust solution to the problem of scheduling quay cranes in container terminals. This method not only improves operational efficiency, but also provides a solid basis for practical applications in port management.

Based on the information provided, here is a diagram representing a hybrid algorithm for solving the Quay Crane Scheduling Problem (QCSP). The steps of components and flow of the hybrid approach, integrating both optimization techniques and heuristic methods.

1. Start: The process begins by initializing the algorithm.
2. Initialize Population: Generate an initial population of feasible schedules for quay cranes.
3. Evaluate Fitness: Assess each schedule based on its makespan and other performance metrics.
4. Selection: Choose the best-performing schedules to create a new generation.
5. Crossover & Mutation: Apply genetic operators to produce new schedules from selected parents.
6. Local Search: Refine these new schedules using local search techniques to improve their quality.
7. Check Termination: Determine if the algorithm has reached a stopping condition, such as a maximum number of iterations or convergence of solutions.
8. Output Best Solution: If termination criteria are met, output the best schedule found; otherwise, return to the evaluation step to continue optimization.

This step illustrates the iterative nature of the hybrid algorithm, emphasizing how it combines genetic algorithms with local search techniques to effectively solve the QCSP in container terminals.

4.1 Case studies and applications

Empirical studies have been carried out to validate these models in real contexts. For example, a study of the EUROGATE terminal in Tangiers enabled a generic model to be adapted to local conditions, demonstrating the effectiveness of mathematical approaches in operational decision-making [2].Here's a sample table that presents hypothetical experimentation results for a hybrid algorithm applied to the Quay Crane Scheduling Problem (QCSP). This table summarizes various metrics that could be used to evaluate the performance of the algorithm across different test scenarios.

Experiment No	Number of cranes	Number of vessels	Makespan(hours)	Total Delays (hours)	Resource utilization	Execution Time (seconds)
1	3	2	12	1	85	45
2	4	3	10	0.5	90	60
3	5	4	15	2	80	75
4	3	5	14	1.5	82	50
5	6	4	11	0.8	88	70
6	4	6	13	1	87	65
7	5	5	12	0.7	89	80
8	3	7	16	2.5	78	55

TAB. 2: Experimentation Results of the Hybrid Algorithm for QCSP

This table provides a clear overview of how the hybrid algorithm performs under varying conditions, highlighting its effectiveness in minimizing makespan and delays while maximizing resource utilization. These metrics are crucial for assessing the algorithm's efficiency and practicality in real-world application

The comparative experimental results between the hybrid algorithm and other optimization methods for the Quayside Crane Planning Problem (QCSP) reveal several significant trends that are worth discussing.

Methods	Makespan (hours)	Totaldelays (hours)	Resource Utilization(%)	execution time (seconds)	Comments
Hybrid algorithm	10	0.5	90	60	best overall performance
pure genetic algorithm	12	1	85	75	less efficient than the hybrid approach
Tabu search	11	0.8	88	70	A good compromise between quality and

					time
ant colony algorithm	13	1.5	82	80	variable performance depending on parameters
method of linear programming	9	0	92	50	optimal for small instances

TAB. 2 : Comparison between methods

4.2Efficiency of the Hybrid Algorithm:

The hybrid algorithm has demonstrated a superior ability to minimise makespan and total delays, achieving optimal performance in complex scenarios. This can be attributed to its ability to combine the advantages of genetic algorithms, which efficiently explore the search space, with local search techniques that refine solutions. This synergy enables more in-depth exploration and efficient exploitation of promising solutions, making the approach robust in the face of operational variations.

4.3. Comparison with other methods:

- Pure genetic algorithm: Although effective, the pure genetic algorithm has shown limitations in terms of delay management. This highlights the need to incorporate local improvement mechanisms to prevent solutions from stagnating in local optima.
- Tabu search: This method provided a good compromise between quality and execution time, but could not compete with the flexibility and adaptability of the hybrid algorithm. Tabu search can sometimes be sensitive to the parameters chosen, which can affect its performance.
- Ant Colony Algorithm: The variable performance of this method suggests that it may require careful adjustment of the parameters for each specific instance of the problem. This may limit its applicability in dynamic environments where conditions change rapidly.
- Linear programming: Although this method gave the best results for small instances, it quickly becomes impractical for more complex scenarios, highlighting the need for more flexible approaches to deal with exponential growth in problem complexity.

Practical implications : The results obtained have significant implications for container terminal managers. The use of a hybrid algorithm could enable more efficient and reactive planning, reducing the costs associated with delays and improving customer satisfaction. In a time-critical environment, this approach could provide a significant competitive advantage.

5 Conclusion and perspectives

Overall, the hybrid algorithm proved to be an effective solution to the Quay Crane Scheduling

Problem (QCSP), outperforming several traditional methods in terms of overall performance. Experimental results demonstrate not only a significant reduction in makespan and delays, but also an optimal use of available resources, which is crucial in a port environment where time is of the essence.

This hybrid approach, which combines the strengths of genetic algorithms with local search techniques, offers remarkable flexibility and adaptability to the dynamic challenges of port operations. The integration of these methods results in solutions that are not only efficient but also robust, able to adapt to unforeseen events such as changes in vessel schedules or equipment failures.

These results pave the way for further research into the optimisation of port operations. It is essential to further explore the potential applications of this hybrid approach in other logistics contexts, including inventory management, transport planning and supply chain optimisation. In addition, the application of machine learning techniques to further refine the algorithms could offer promising prospects for further improving operational performance.

References

[1] M. Gulić et M. Žuškin, « Enhancing Metaheuristic Optimization: A Novel Nature-Inspired Hybrid Approach Incorporating Selected Pseudorandom Number Generators », *Algorithms*, vol. 16, n° 9, p. 413, août 2023, doi: 10.3390/a16090413.

[2] J. H. Chen, D.-H. Lee, et M. Goh, « An effective mathematical formulation for the unidirectional cluster-based quay crane scheduling problem », *European Journal of Operational Research*, vol. 232, n° 1, p. 198-208, janv. 2014, doi: 10.1016/j.ejor.2013.06.051.

[3] C. Lersteau et W. Shen, « A survey of optimization methods for Block Relocation and PreMarshalling Problems », *Comput. Ind. Eng.*, vol. 172, p. 108529, 2022, doi: 10.1016/j.cie.2022.108529.

[4] A. Agra et M. Oliveira, « MIP approaches for the integrated berth allocation and quay crane assignment and scheduling problem », *European Journal of Operational Research*, vol. 264, n° 1, p. 138-148, janv. 2018, doi: 10.1016/j.ejor.2017.05.040.

[5] S. Aidi et M. Mazouzi, « Optimization Approach for Yard Crane Scheduling Problem using genetic algorithm in Container Terminals », *ITM Web Conf.*, vol. 52, p. 02002, 2023, doi: 10.1051/itmconf/20235202002.

[6] A. Skaf, S. Lamrous, Z. Hammoudan, et M. A. Manier, « Single quay crane and multiple yard trucks scheduling problem with integration of reach-stacker cranes at port of Tripoli-Lebanon », in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy: IEEE, oct. 2019, p. 852-857. doi: 10.1109/SMC.2019.8914667.

[7] D. Feillet, S. N. Parragh, et F. Tricoire, « A Local-Search Based Heuristic for the Unrestricted Block Relocation Problem », *Comput. Oper. Res.*, vol. 108, p. 44-56, 2018, doi: 10.1016/j.cor.2019.04.006.

[8] S. Aidi, I. Torbi, M. Mazouzi, et I. Belassiria, « Planning of one-way quay cranes using the genetic algorithm », *SHS Web Conf.*, vol. 119, p. 07005, 2021, doi:

10.1051/shsconf/202111907005.

[9] O. Karpuzoglu, M. Akyüz, et T. Öncan, « A Tabu Search Based Heuristic Approach for the Dynamic Container Relocation Problem », 2015, doi: 10.1007/978-3-319-42902-1_22.

[10] K. Chargui, T. Zouadi, V. Sreedharan, A. E. Fallahi, et M. Reghioui, « A novel robust exact decomposition algorithm for berth and quay crane allocation and scheduling problem considering uncertainty and energy efficiency », *Omega*, vol. null, p. null, 2023, doi: 10.1016/j.omega.2023.102868.

[11] A. Skaf, S. Lamrous, Z. Hammoudan, et M.-A. Manier, « Integrated quay crane and yard truck scheduling problem at port of Tripoli-Lebanon », *Computers & Industrial Engineering*, vol. 159, p. 107448, sept. 2021, doi: 10.1016/j.cie.2021.107448.

[12] T. Bacci, S. Mattia, et P. Ventura, « A branch-and-cut algorithm for the restricted Block Relocation Problem », *Eur. J. Oper. Res.*, vol. 287, p. 452-459, 2020, doi: 10.1016/j.ejor.2020.05.029.

[13] A. Sanaa, T. Imane, et M. Mohamed, « Quay Crane Scheduling in Container Terminals Using a Hybrid Genetic Algorithm », *Journal of Computer Science*, 2025.

[14] H. Amani, L. Bouyaya, R. Chaib, F. Z. Djekrif, et M. Aizi, « Optimisation of Quay Crane Scheduling Problem at the Port of Algeria », in *Reliability and Statistics in Transportation and Communication*, vol. 640, I. Kabashkin, I. Yatskiv, et O. Prentkovskis, Éd., in Lecture Notes in Networks and Systems, vol. 640. , Cham: Springer International Publishing, 2023, p. 232-241. doi: 10.1007/978-3-031-26655-3_21.

A multi-stage stochastic model for the agri-food supply chain with labor constraints and blockchain technology

Clément Bamogo[1,2] Salma Garra[1] Mustapha Oudani[3]
Amine Belhadi[1] Sofiene Dellagi[2] Zied Achour[2]

[1] BearLab, Rabat Business School
[2] LGIPM, Lorraine University
[3] TICLAB, College of Engineering and Architecture

Abstract

This paper examines the specifics of forced labor in the agri-food sector. This paper examines the agricultural supply chain, identifying the main stakeholders, including farmers, buyers, recruiters, etc., and describing the flows of goods and labor within this chain. Indeed, labor-intensive operations such as farming and harvesting can involve high costs for small farmers and a risk of contracting illegal labor such as forced labor. To mitigate these issues, the paper proposes a multi-stage stochastic model that incorporates Blockchain technology to ensure transparency and accountability throughout the supply chain. The proposed approach aims to maximize overall profit and minimize costs while having good transparency on harmful labor practices. By leveraging Blockchain technology, stakeholders can validate and accept transactions, ensuring fair labor practices and ethical operations in the supply chain. This paper sets the scene for an in-depth study that combines theoretical ideas with practical solutions to urgent problems in the agri-food supply chain.

Keywords : *Forced labor, Agri-food, Supply chain, Small farmer, Blockchain technology, Stochastic modeling.*

1 Introduction

Effective logistics planning is an integral part of supply chain management (SCM), it is defined as complex networks of interconnected entities. This network can be very complex because of the multitude of goods and services with which companies are frequently involved, as is the case in agricultural supply chains, where there may be farmers, buyers, pre-processors, processors, manufacturers, recruiters, etc., all involved in the supply chain (Young and Esqueda, 2005). The supply chain goes beyond individual actors to become a multi-level chain or multi-echelon supply chain as described by Manupati et al. (2020). The network encompasses all entities involved in the value-adding or transformation process of raw materials or services into finished products. These products are then delivered to the end customer to meet specific needs, involving the management of materials, information, and financial resources throughout the process. Supply chain management is then defined as the optimal management of these relationships as described by (Rivera-Cadavid et al., 2022). Research into mathematical optimization for profit maximization and cost minimization under constraints exists for specific parts of the agri-food supply chain, including those that take labor management into account. Labor is a vital resource in supply chains, from planting through transport and storage to distribution, as underlined by Agyemang et al. (2022). In their research, Allain et al. (2013) highlight the notion of *"labor supply chain"* in global supply chain network as a major traceability challenge in today global supply chain. In fact, the author highlights the existence of intermediaries in the recruitment process, whose transparency and sources of labor remain uncertain, leading to a risk of forced labor for workers. Some researchers and workers' rights organizations have already mentioned the importance of ethical considerations in the agricultural supply chain, while pointing the finger at the origins or sources of recruitment. The International Labor Organization (ILO) and Walk Free supported by the work of some researchers have highlighted the risks associated with illegal sources of recruitment process of workers who may find themselves in situations of forced labor (Benstead et al., 2020). Forced labor is recognized as a persistent and ingrained aspect of the modern global economy (LeBaron and Gore, 2020). Policymakers and civil society organizations worldwide are investing significant efforts and resources to combat this issue. This situation is more alarming in Africa, where a large part of the population lives in agriculture and depends heavily on international trade. Several issues, such as political instability, poverty, and corruption, make Africa the most vulnerable region to modern slavery in the world, with the highest vulnerability at 64%, with 3.8 million people in forced labor on the continent (Free, 2023). These statistics hide disparities due to difficulties in tracing the recruitment process and attempts to understand forced labor patterns and operations in global supply chains have been hampered by the lack of in-depth, impartial, and comprehensive empirical data (LeBaron and Gore, 2020). This influence has surged, coinciding with the emergence of neo-liberal trends like privatization, fragmented regulatory oversight, and globalization (Davies, 2020). As a result, the prevalence of multinational corporations has intensified price pressures on small farmers in the supply network. This escalation has augmented both supply and demand pressures, often leading farmers to the adoption of subcontracted, flexible labor practices, as noted by the EHRC (2010). In fact, to meet their profit, downstream farmers can carry out illegal recruitment processes to satisfy demand and ensure a substantial gain. These market dynamics create opportunities for unethical employers and intermediaries to exploit their employees (Davies, 2020) and this point of view is supported by the works of Tombs and Whyte (2015) when they consider supply chain network as a possible *"technique for contracting out crime"*. In addition, seasonal work and the prevalence of undocumented migrants can be ways of employing workers with forced labor characteristics (Davies, 2020). As undocumented migrants face problems related to detention and deportation in the event of denunciation by their employers, they end up working under inappropriate and even threatening conditions, notes the EHRC (2010). Labor rights concerns within the agricultural sector garner substantial attention from various stakeholders, highlighting the importance of addressing social sustainability issues. In developing countries, small and medium-scale enterprises (SMEs) play a pivotal role in the sector's supply chain, making their involvement essential for implementing initiatives aimed at tackling these challenges (Agyemang et al., 2022). Despite the existence of traditional means such as audits to regulate recruitment in agricultural supply chains, it is clear that they are ineffective. A social audit involves examining labor standards within a company through activities like factory visits, document analysis, and interviews with both workers and managers. It aims to assess various aspects of working conditions such as working hours, harassment policies, health and safety measures, and the prevention of child labor (Anisul Huq et al., 2014). New (2015) considers the audit to be a *"visit pre-notified"* and therefore without effect on the fight against forced labor in the agri-food supply chain, or is even pointless. What's more, audits focus on a specific point in the network and are therefore not very effective when it comes to looking at the whole SCM, which is an extensive complex network (Benstead et al., 2020). Even with access to stakeholder labor recruitment histories on databases, another challenge of audits is corruption, in the sense that companies or farmers visited can obtain certification by paying a bribe to representatives of the government or approved auditors. Indeed, if suppliers can circumvent certain aspects of codes of conduct, it is partly because corruption is commonplace (Anisul Huq et al., 2014).

Blockchain technology has emerged as a pivotal force in the transformation of supply chains, owing to its intrinsic characteristics like decentralization, transparency, and immutability. Furthermore, it facilitates secure commercial transactions and increases confidence between entities through the utilization of smart contracts (Shahid et al., 2020). Blockchain's effectiveness in addressing traceability challenges stems from its decentralized, distributed, and immutable information structure. This architecture enhances resilience against counterfeit certificates and provides a sophisticated system for achieving end-to-end traceability in the supply chain as explained by Biswas et al. (2023). The strength of distributed system technologies is leading various industries around the world to recognize the emergence of blockchain technology, including finance, electronic medical records (EMR), the internet of things (IoT), energy, and many others add Shahid et al. (2020). Our paper aims to contribute to the growing work on blockchain-based agri-food supply chains and provide a multi-stage stochastic linear model solution integrating labor management across the whole agri-food supply chain from farming tasks to manufacturers through pre-processing stages. The work is organized as follows: in section 2, we will develop the multi-stage stochastic model of the problem. In section 3 we propose a small numerical example and analyze the result obtained from the model. Section 4 is reserved for managerial insights and the paper ends with a conclusion in section 5, which includes prospects for extending the work.

2 Problem formulation and mathematical model

2.1 Problem description

We consider a network consisting of a set F representing the farmers, a set R for recruiters, and a set B for product buyers (intermediaries, focal firms, collectors, ...). We assume that we should study the supply chain of a single product between farmers and buyers and the recruitment flow between recruiters and farmers. Farmers cultivate abundant crops and meticulously oversee their growth, ensuring quality and monitoring progress. Subsequently, they supply these crops downstream. To meet his production at a period, each farmer has to recruit enough workers from recruiters.

Let's consider the tables Tab. 1 and Tab. 2 that contain the sets, the data, and decision variables of the model:

TAB. 1: Sets used by the model.

Sets	**Index**	**Description**
$R = \{1, 2, \cdots, r\}$	l	Set of recruiters
$F = \{1, 2, \cdots, n\}$	i	Set of farmers
$B = \{1, 2, \cdots, m\}$	j	Set of buyers
$T = \{1, 2, \cdots, d\}$	t	Set of periods
$K = \{1, 2, \cdots, s\}$	k	Set of scenarios

TAB. 2: Data and decision variables.

Variable	**Description**
$Y_{i,t}^k$	Goods production of farmer i during period t in scenario k
$I_{i,t}^k$	Inventory of stock for farmer i at the end of period t in scenario k
$\alpha_{i,t}^k$	Level of good practice of recruitment by farmer i at period t in scenario k
$\Delta_{i,t}^k$	Amount of money that the farmer i invests in blockchain at period t in scenario k
$L_{i,l,t}^k$	1 if farmer i can do external recruitment from recruiter l at period t, 0 else in scenario k
$\Gamma_{i,l,t}^k$	External number of employees provided by recruiter l for farmer i at period t in scenario k
$X_{i,j,t}^k$	Quantity of goods farmer i delivered to buyers j at period t in scenario k
Data	**Description**
ρ	Sensitivity of buyers and recruiters to the traceability of the workers
$\delta_{j,t}^k$	Quantity of goods demand for buyers j at period t for scenario k

Pr_k	Probability of scenario of k
$p_{j,t}$	price proposed by buyer j for the targeted product at period t
$\gamma^k_{i,l,t}$	Operational salary of labor payed by farmer i from recruiter l for scenario k
$\psi_{l,t}$	Available number of employees in farming tasks with the recruiter l at period t
$\sigma_{i,t}$	Number of workers required by farmer i to produce a unit of product at period t
Q_i	Capacity of production for farmer i
$B_{i,t}$	Labor budget for farmers i at period t
$c_{i,j,t}$	Operational unit cost of transportation between i to intermediary j at period t
$\lambda_{i,t}$	unit cost of inventory management for farmer i at period t
U	Amount of money that must pay a farmer every period for the social audit

We now propose a multi-stage stochastic programming approach to incorporate the uncertainty associated with demands, supplies, processing costs, transportation costs, shortage costs and capacity expansion costs. This approach is proposed in order to take into account the random phenomena that can impact forecasts, such as in this case buyer demand. In fact, despite its simplicity the fundamental assumption of deterministic programming, wherein the problem entries are assumed to be known fixed data, frequently does not align with reality (Mayer and Kall, 2011). Instead, it is common for (some of) the entries to be derived from statistical estimates based on observed real data, such as samples. Alternatively, in certain cases, the model design may dictate that the entries represent random variables, such as capacities, demands, productivities, or prices. Modeling a multi-stage stochastic problem involves decision-making over multiple periods under uncertainty. The decisions at each stage depend on the outcomes realized in previous stages, and uncertainty is typically modeled using a set of scenarios k where Pr_k represent the probability of the scenario k and $\sum_{k=1}^{s} Pr_k = 1$ as described by Mayer and Kall (2011) and Shapiro et al. (2021). Then it makes sense to talk about the expected objective value $\mathbb{E}[Z]$ instead of the objective value Z because we calculate the expected value considering the set of scenarios k.

We first describe a stochastic formulation for the problem of agri-food supply chain without blockchain where we consider the distribution to be known.

$$\left.\begin{array}{ll} \text{Max } \mathbb{E}[Z] & = \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \sum_{j=1}^m (p_{j,t} - c_{i,j,t}) X^k_{i,j,t}) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \sum_{l=1}^r \gamma^k_{i,l,t}\Gamma^k_{i,l,t}) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \lambda_{i,t}(Y^k_{i,t} + I^k_{i,t-1})) \end{array}\right\} \tag{1}$$

Subject to :

$$Y^k_{i,t} \leq Q_i - I^k_{i,t-1} \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{2}$$

$$\sum_{i=1}^{n} X^k_{i,j,t} = \delta^k_{j,t} \quad \forall j \in C, \quad \forall t \in T, \quad \forall k \in K \tag{3}$$

$$\sum_{j=1}^{m} X^k_{i,j,t} \leq I^k_{i,t-1} + Y^k_{i,t} \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{4}$$

$$I^k_{i,t} = I^k_{i,t-1} + Y^k_{i,t} - \sum_{j=1}^{m} X^k_{i,j,t} \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{5}$$

$$\sum_{l=1}^{r} \Gamma^k_{i,l,t} \geq \sigma_{i,t} Y^k_{i,t} \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{6}$$

$$\sum_{i=1}^{n} \Gamma^k_{i,l,t} \leq \psi_{l,t} \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{7}$$

$$X_{i,j,t}^k \geq 0 \quad \forall i \in F, \quad \forall j \in C, \quad \forall t \in T, \quad \forall k \in K \tag{8}$$

$$Y_{i,t}^k \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{9}$$

$$I_{i,t}^k \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{10}$$

$$\Gamma_{i,l,t}^k \in \mathbb{N} \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{11}$$

In the above model, Z denotes the expected value of the total profit for the set of farmers. The objective function (1) includes farmers income $(p_{j,t} - c_{i,j,t})X_{i,j,t}^k$, the total cost of labor $\sum_{l=1}^{r} \gamma_{i,l,t}^k \Gamma_{i,l,t}^k$, the inventory cost $\lambda_{i,t}(Y_{i,t}^k + I_{i,t-1}^k)$. The constraint (3) implies that every buyer is satisfied in every period. Constraint (2) states that a farmer cannot produce more than available space in his field and constraint (4) means that a farmer cannot deliver products more than available at period t. Constraint (5) represents the update of inventory level for each farmer. Depending on the quantity of production $Y_{i,t}^k$, each farmer can do external recruitment as shown in constraint (6) to meet his needs. Constraints (7) represent respectively the maximum available number of workers . Constraints (8) to (11) are non-negativity constraints for decision variables.

2.2 Multi-stage stochastic model for blockchain-based supply chain

In a blockchain-based agri-food supply chain, three additional decision variables are added to the traditional SC design problem. The first variable $L_{i,l,t}^k \in \{0,1\}$ represents the decision of farmer i to recruit workers from recruiter l at period t, this means that $L_{i,l,t}^k = 1$ if farmer i can recruit worker from l at period t and 0 else. The second decision variable is $\alpha_{i,t}^k \in [0,1]$ to represent the level of good practice of farmer i l at period t. The third variable is the investment $\Delta_{i,t}^k \geq 0$, the amount of money that farmers invest each period to keep the transparency and the level of good practice high. Consider the stochastic model below, integrating blockchain into the agricultural supply chain:

$$\left.\begin{aligned} \text{Max } \mathbb{E}[Z] \quad &= \textstyle\sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \sum_{j=1}^m (p_{j,t} - c_{i,j,t})X_{i,j,t}^k) \\ &- \textstyle\sum_k^s Pr_k(\sum_{t=1}^d \sum_{j=1}^m \sum_{l=1}^r \gamma_{i,l,t}^k (1+\alpha_{i,t}^k)\Gamma_{i,l,t}^k) \\ &- \textstyle\sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \lambda_{i,t}(Y_{i,t}^k + I_{i,t-1}^k)) \\ &- \textstyle\sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n (1-\alpha_{i,t}^k)U) \\ &- \textstyle\sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \Delta_{i,t}^k) \end{aligned}\right\} \tag{12}$$

Subject to :

(3) − (4) − (5) − (6) − (7)

$$Y_{i,t}^k \leq (1+\alpha_{i,t}^k)Q_i - I_{i,t-1}^k \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{13}$$

$$\Gamma_{i,l,t}^k \leq \psi_l \cdot L_{i,l,t}^k \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{14}$$

$$L_{i,l,t}^k \leq \frac{\alpha_{i,t}^k}{\pi_t} \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{15}$$

$$\alpha_{i,t}^k = \rho\Delta_{i,t}^k + \alpha_{i,t-1}^k \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{16}$$

$$X_{i,j,t}^k \geq 0 \quad \forall i \in F, \quad \forall j \in C, \quad \forall t \in T, \quad \forall k \in K \tag{17}$$

$$Y_{i,t}^k \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{18}$$

$$I_{i,t}^k \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{19}$$

$$\Delta_{i,t}^k \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{20}$$

$$\alpha_{i,t}^k \in [0,1] \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{21}$$

$$\Gamma_{i,l,t}^k \in \mathbb{N} \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{22}$$

$$L^k_{i,l,t} \in \{0,1\} \quad \forall i \in F, \quad l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{23}$$

In the objective function (12), $(1+\alpha^k_{i,t})$ represents the increase in traceability and the level of good practices regarding working conditions in the farmer i activities in scenario k compared to traditional SC. The quantity of products $X^k_{i,j,t}$ sold by a farmer i to a buyer j will depend on the level of good practice $\alpha^k_{i,t}$. The higher the $\alpha^k_{i,t}$ level, the more the farmer can increase his production capacity as shown in constraint (13). However, a higher level of good recruitment processes implies a higher cost in the wages of external workers and needs more investment until the level comes to maximum ($\alpha^k_{i,t} = 1$). In addition, contrary to the traditional supply chain scenario, social auditing costs will decrease when the good practice level $\alpha^k_{i,t}$ is getting high and we go from U in traditional SC to $(1-\alpha^k_{i,t})U$ in blockchain-based one. In the constraint (15) the boolean $L^k_{i,l,t}$ can be 1 if and only if $\alpha^k_{i,t} \geq \pi_t$. In other words, for a farmer i to be able to recruit external workers, he must have a level $\alpha^k_{i,t}$ that reaches at least the minimum approval rate π_t. The constraint (16) represents a condition of updating the approval rate $\alpha^k_{i,t}$ that depends on the investment $\Delta^k_{i,t}$ and the previous rate $\alpha^k_{i,t-1}$.

2.3 Linear multi-stage stochastic model

We propose to linearize the objective function (24) by introducing the decision variable $W_{i,l,t}$ such that $W_{i,l,t} = \alpha^k_{i,l}\Gamma^k_{i,l,t}$. The variables $W_{i,l,t}$ is under the following conditions:

$$\left.\begin{array}{l} W_{i,l,t} \leq \Gamma^k_{i,l,t} \\ \rho\Delta^k_{i,t} \leq \frac{\sum_{l=1}^{r} W_{i,l,t}}{M} \end{array}\right\}$$

The model becomes then :

$$\left.\begin{array}{ll} \text{Max } \mathbb{E}[Z] & = \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \sum_{j=1}^m (p_{j,t} - c_{i,j,t})X^k_{i,j,t}) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{j=1}^m \sum_{l=1}^r \gamma^k_{i,l,t}(\Gamma^k_{i,l,t} + W_{i,l,t})) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \lambda_{i,t}(Y^k_{i,t} + I^k_{i,t-1})) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n (1-\alpha^k_{i,t})U) \\ & - \sum_k^s Pr_k(\sum_{t=1}^d \sum_{i=1}^n \Delta^k_{i,t}) \end{array}\right\} \tag{24}$$

Subject to :

(13) – (14) – (15) – (16)

$$W_{i,l,t} \leq \Gamma^k_{i,l,t} \tag{25}$$

$$\rho\Delta^k_{i,t} \leq \frac{\sum_{l=1}^{r} W_{i,l,t}}{M} \tag{26}$$

$$X^k_{i,j,t} \geq 0 \quad \forall i \in F, \quad \forall j \in C, \quad \forall t \in T, \quad \forall k \in K \tag{27}$$

$$Y^k_{i,t} \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{28}$$

$$I^k_{i,t} \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{29}$$

$$\Delta^k_{i,t} \geq 0 \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{30}$$

$$\alpha^k_{i,t} \in [0,1] \quad \forall i \in F, \quad \forall t \in T, \quad \forall k \in K \tag{31}$$

$$\Gamma^k_{i,l,t} \in \mathbb{N} \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{32}$$

$$W_{i,l,t} \in \mathbb{N} \quad \forall i \in F, \quad \forall l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{33}$$

$$L^k_{i,l,t} \in \{0,1\} \quad \forall i \in F, \quad l \in R, \quad \forall t \in T, \quad \forall k \in K \tag{34}$$

3 Numerical example

Let's consider an agricultural chain consisting of 3 farmers, 2 buyers, and 3 external recruitment sources. These groups will have to carry out their activities over 12 periods. Considering the farmers as the epicenter of this chain, they have 4 scenarios with distinct probabilities on the possibilities of orders from the buyers. We illustrate this supply chain as the figure 1 below :

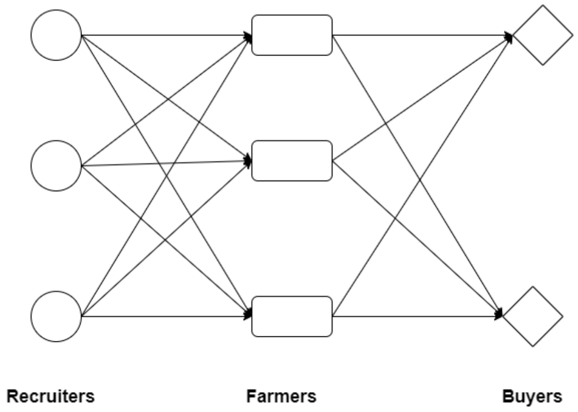


FIG. 1: Supply chain design problem

The sets are summarized in Tab. 3 and scenarios in Tab. 4 below:

TAB. 3: Set of the Supply chain network.

Parameters	Values
Farmers	F1, F2, F3
Recruiters	R1, R2, R3
buyers	B1, B2
Periods	10
Scenarios	4

TAB. 4: Set of scenarios.

Scenario	S1	S2	S3	S4
Probability	0.50	0.25	0.18	0.07

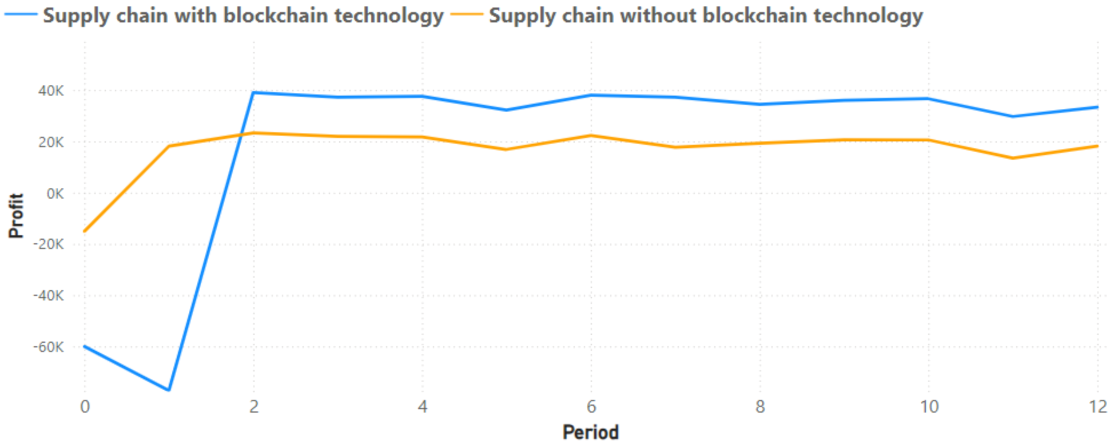


FIG. 2: Profit over periods with and without blockchain technology

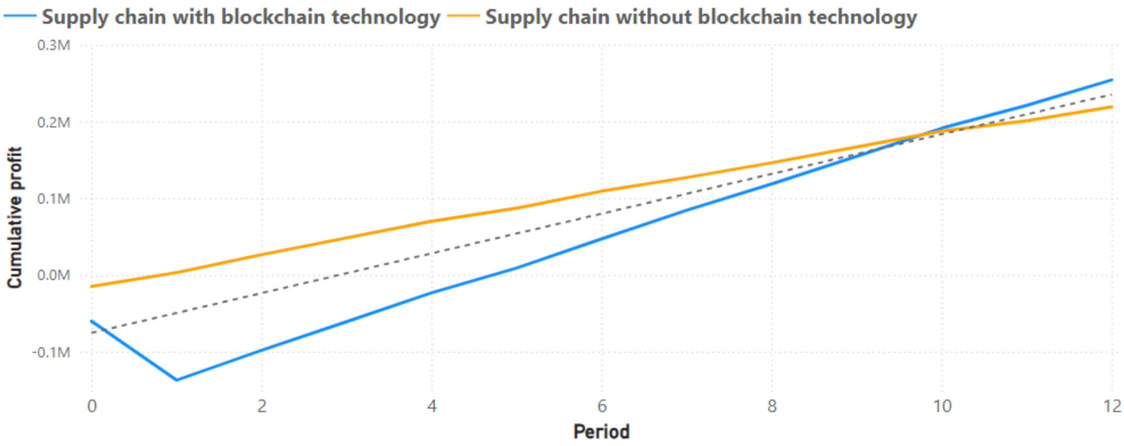


FIG. 3: Cumulative profit evolution

Figure 2 and 3 show the evolution of overall costs over the various periods. The periods $t = 0$ and $t = 1$ were negatively balanced due to the costs of integration and investment in the blockchain, as shown in figure 2. However, over time (from $t = 2$), we observe higher profits in the system using blockchain. Figure 3 confirms these results, with highly optimal profits over the long term and a rising trend line. It shows the cumulative profit evolution in both systems, where profits in the blockchain-based system remaining low until a number of periods before surpassing the traditional case.

An analysis of profits according to scenario provides some interesting information, as shown in figure 4. In fact, the observations show fairly high profits compared with traditional AFSCs when the probability of the scenarios occurring is high. This would mean that the profitability of blockchain would have to be based on high product order certainty.

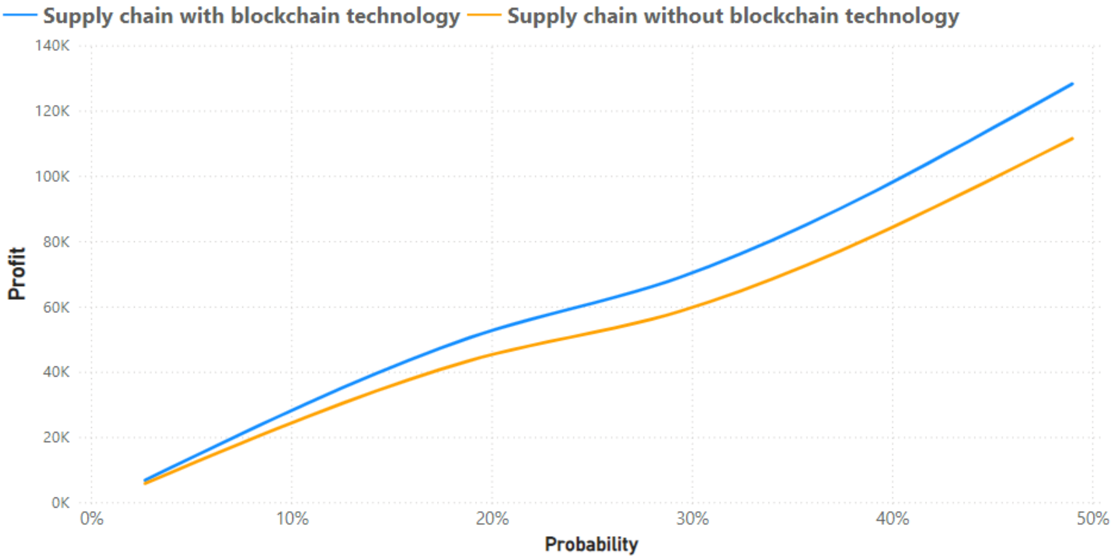


FIG. 4: Profit evolution following probabilities of scenarios

4 Managerial insights

Our study offers several critical managerial insights based on the observation and interpretation of supply chain evolution over different periods and scenarios. These insights are vital for managers aiming to optimize the agri-food supply chain and combat forced labor effectively.

- **Insight 1:** Initial investment and long-term cost benefits
  The integration of blockchain technology into the agri-food supply chain incurs a sig-

nificant initial cost, primarily due to the need for farmers and other stakeholders to ensure compliance with ethical practices before transactions can occur via blockchain. As illustrated in Figure 2 and Figure 3, the initial periods exhibit higher costs attributable to these investments. However, managers should recognize that this initial expenditure is an investment in long-term cost reduction. Over subsequent periods, systems utilizing blockchain demonstrate a consistent decrease in overall costs, unlike traditional systems where costs remain high.

- **Insight 2:** Stochastic-based management
  The analysis of supply chain behavior across different stochastic scenarios, as depicted in Figures 4 provides valuable insights for managing costs and risks. Traditional supply chains consistently exhibit high forced labor risks across all scenarios. In contrast, blockchain systems can maintain lower forced labor risks if the scenarios with high probabilities of occurrence are realized. However, the overall costs of blockchain systems can fluctuate depending on these probabilities.

  Managers could, therefore, adopt a proactive approach to scenario planning, ensuring that high-probability scenarios are prioritized to keep costs manageable while minimizing forced labor risks. This involves continuous monitoring of market conditions, labor practices, and potential disruptions to anticipate and respond to changes effectively.

In summary, the integration of blockchain technology into the agri-food supply chain offers substantial long-term benefits in terms of cost reduction and risk mitigation of forced labor. Managers may navigate the initial investment phase strategically, prioritize high-probability scenarios, continuously monitor and adapt to changes, and foster collaboration and capacity building to ensure sustainable and ethical supply chain operations.

5 Conclusion and perspectives

Determining the optimal SC configuration is a difficult problem since a lot of factors and objectives must be taken into account when designing the network under uncertainty. The agri-food supply chain faces multifaceted challenges, from fluctuating demand and supply uncertainties to labor exploitation and market demand. Our paper proposed a model to optimize the product and labor supply chain and a framework for a blockchain-based supply chain for good practices concerning some illegal practices in labor recruitment. Using a stochastic approach, we have identified the problem of uncertainty and analyzed the results according to different scenarios. Decision-makers should have good predictive methods and a long enough track record to handle these uncertainties. However, all the results observed show that a supply chain based on blockchain could well meet a social interest by increasing transparency in recruitment, while at the same time being economically beneficial to small-scale farmers in the long term. For future work, this problem could be extended into a multi-objective problem to emphasize the concept of illegal or forced labor risk. Using a Pareto optimum analysis, we can simulate the risk level of these harms and the profits generated in the two agricultural supply chain systems.

References

Agyemang, M., Kusi-Sarpong, S., Agyemang, J., Jia, F., and Adzanyo, M. (2022). Determining and evaluating socially sustainable supply chain criteria in agri-sector of developing countries: insights from west africa cashew industry. *Production Planning & Control*, 33(11):1115–1133.

Allain, J., Crane, A., LeBaron, G., and Behbahani, L. (2013). Report: Forced labour's business models and supply chains.

Anisul Huq, F., Stevenson, M., and Zorzini, M. (2014). Social sustainability in developing country suppliers: An exploratory study in the ready made garments industry of bangladesh. *International Journal of Operations & Production Management*, 34(5):610–638.

Benstead, A. V., Hendry, L. C., and Stevenson, M. (2020). Detecting and remediating modern slavery in supply chains: a targeted audit approach. *Production Planning & Control*, 32(13):1136–1157.

Biswas, D., Jalali, H., Ansaripoor, A. H., and De Giovanni, P. (2023). Traceability vs. sustainability in supply chains: The implications of blockchain. *European Journal of Operational Research*, 305(1):128–147.

Davies, J. (2020). Corporate harm and embedded labour exploitation in agri-food supply networks. *European Journal of Criminology*, 17(1):70–85.

EHRC (2010). Inquiry into recruitment and employment in the meat and poultry processing sector.

Free, W. (2023). Modern slavery in africa.

LeBaron, G. and Gore, E. (2020). Gender and forced labour: Understanding the links in global cocoa supply chains. *The journal of development studies*, 56(6):1095–1117.

Manupati, V. K., Schoenherr, T., Ramkumar, M., Wagner, S. M., Pabba, S. K., and Inder Raj Singh, R. (2020). A blockchain-based approach for a multi-echelon sustainable supply chain. *International Journal of Production Research*, 58(7):2222–2241.

Mayer, J. and Kall, P. (2011). Stochastic linear programming. models, theory, and computation. *International series in operations research & management science*, 156.

New, S. J. (2015). Modern slavery and the supply chain: the limits of corporate social responsibility? *Supply Chain Management: An International Journal*, 20(6):697–707.

Rivera-Cadavid, L., Peña-Orozco, D. L., Bravo-Bastidas, J. J., Puertas-Marín, A. F., and Falla-Ramírez, Y. T. (2022). Harvest and distribution planning model for a fruit supply chain. *Revista UIS Ingenierías*, 21(2):115–130.

Shahid, A., Almogren, A., Javaid, N., Al-Zahrani, F. A., Zuair, M., and Alam, M. (2020). Blockchain-based agri-food supply chain: A complete solution. *Ieee Access*, 8:69230–69243.

Shapiro, A., Dentcheva, D., and Ruszczynski, A. (2021). *Lectures on stochastic programming: modeling and theory*. SIAM.

Tombs, S. and Whyte, D. (2015). *The corporate criminal: Why corporations must be abolished*. Routledge.

Young, R. R. and Esqueda, P. (2005). Vulnerabilidades de la cadena de suministros: consideraciones para el caso de américa latina. *Academia. Revista latinoamericana de administración*, (34):63–78.

AI-Driven Data Management and Analytics for Industry 4.0: Optimizing Supply Chains and Enhancing Safety in Mining

BIADI Wafaa[1], EL BOUZEKRI EL IDRISSI Adiba[1], ASSIF Safaa [1], HANI Yasmina[2]
[1] Chouaïb Doukkali University, Natl Schl Appl Sci, Sci Engineer Lab Energy
El Jadida, Morocco
(Email: biadi.wafaa@ucd.ac.ma)
(Email : elbouzekrielidrissi.a@ucd.ac.ma)
(Email:assif.s@ucd.ac.ma)
[2] Paris 8 University, Lab QUARTZ
Paris, France
(Email: y.hani@iut.univ-paris8.fr)

Abstract:

Ensuring workforce safety while optimizing supply chain operations in the mining industry, within the context of Industry 4.0, presents significant challenges. Leveraging advanced technologies such as artificial intelligence (AI), real-time data management, and analytics offers innovative solutions to address these complexities. This article explores how AI-driven data management and analytics can optimize supply chains and enhance safety protocols in mining operations.

In the era of Industry 4.0, AI enables real-time monitoring of mining environments, predicting equipment failures, and assessing potential safety hazards before they become critical. By analyzing vast datasets, AI can uncover patterns and trends that would otherwise go unnoticed, leading to proactive decision-making in safety management. Furthermore, advanced analytics can optimize supply chain processes, ensuring more efficient resource allocation, reducing downtime, and improving overall operational efficiency.

Data visualization techniques further enhance these insights, providing intuitive representations of complex data, and enabling stakeholders to make informed, timely decisions. The integration of these technologies not only streamlines supply chain logistics but also plays a crucial role in risk mitigation, accident prevention, and the continuous improvement of working conditions in the mining sector.

This paper discusses the potential of AI-driven systems, a core element of Industry 4.0, to transform both safety and supply chain efficiency in mining. It offers practical applications and case studies where these technologies have already made a measurable impact. By combining the power of AI, real-time monitoring, and data analytics, mining companies can enhance operational safety while optimizing supply chain performance, creating a more sustainable and secure working environment in line with Industry 4.0 advancements.

Keywords: *Mining industry; Industry 4.0; Supply chain optimization; safety; Artificial intelligence (AI); Real-time monitoring; data management; Advanced analytics.*

1 Introduction:

The words "industry 4.0 (I4.0)" and "fourth industrial revolution (IR4.0)" were invented to denote the significant advancements in automation and machine intelligence. By conflating digital, physical, and biological phenomena, IR 4.0 represents the relationship between humans and technology **[1].** Businesses must now adopt terminology such as advanced manufacturing, augmented reality, analytical systems, virtual simulations, digital production management, and industrial robotics. This requires innovation, which is described as a series of actions that result in the development and initial application of novel technical, procedural, or organizational solutions **[2].**

The convergence of these technologies in Industry 4.0 not only enhances automation but also transforms the role of human labor, fostering collaboration between humans and machines. This shift enables more efficient and safer working environments, where real-time data and advanced analytics play a pivotal role in decision-making processes. As machines become more intelligent and autonomous, industries must rethink their traditional approaches, integrating innovative technologies such as AI, the Internet of Things (IoT), and smart systems to create more responsive, adaptable, and interconnected operations. This evolution is vital for businesses looking to harness the full potential of IR 4.0.

The mining sector, much like other industries, faces a range of inherent complexities that stem from its demanding environments, sophisticated processes, and vast supply chains **[3].** These challenges require continuous adaptation and innovation to ensure that operations run smoothly and safely. The integration of Industry 4.0 technologies has emerged as a critical solution for addressing these complexities **[4].**

In a sector where operations generate massive amounts of data—whether it's from equipment performance, environmental factors, or logistical processes—the ability to manage and analyze this data effectively becomes crucial **[5]**. Industry 4.0 technologies provide the tools to do just that, allowing companies to optimize their performance, enhance safety measures, and improve overall productivity **[6]**.

Moreover, as industries worldwide evolve toward smarter, more interconnected systems, the adoption of these advanced technologies has moved beyond being a competitive advantage to becoming a necessity **[7]**. In today's rapidly changing industrial landscape, remaining relevant and efficient requires embracing innovations that can transform traditional operations into more agile, responsive systems. Technologies such as AI-driven analytics, IoT-based monitoring, and automation are no longer optional add-ons; they are essential for companies aiming to thrive in an increasingly digital and data-driven world **[8].**

These tools are reshaping industries by driving improvements in efficiency, elevating safety standards, and offering new levels of flexibility and adaptability to address unforeseen challenges **[9]**.

In this context, the mining industry must embrace these technologies to remain competitive and navigate its unique challenges **[10]**. By doing so, it positions itself not only to optimize day-to-day operations but also to stay at the forefront of the broader shift toward more sustainable, connected, and intelligent industrial ecosystems. The transformation enabled by Industry 4.0 will be key in determining the future trajectory of industries like mining, as companies strive to balance operational demands with the need for greater efficiency, safety, and responsiveness **[11]**.

2 The key pillars of Industry 4.0 in mining

In an era of rapid technological advancement, Industry 4.0 represents a significant leap in innovation and transformation, especially within the mining sector. By merging digital technologies with traditional industrial processes, it brings unprecedented productivity, connectivity, and efficiency. At its core, Industry 4.0 integrates key technologies such as the Internet of Things (IoT), Big Data analytics, Artificial Intelligence (AI), and automation (Figure 1), creating a connected and intelligent ecosystem. These technologies enable real-time monitoring, predictive maintenance, and data-driven decision-making, leading to safer, more efficient, and sustainable mining operations. Together, these pillars form a cohesive framework that optimizes production, reduces costs, and enhances adaptability across the industry.

2.1 IoT (Internet of Things) and Sensors

The Internet of Things (IoT) refers to a network of interconnected devices that communicate and exchange data over the Internet. In the context of Industry 4.0, IoT sensors play a pivotal role by enabling real-time monitoring and data collection across various stages of the manufacturing process. These sensors facilitate data-driven decision-making by providing critical insights into machine performance, product quality, and energy consumption. Through continuous data acquisition, IoT technology supports predictive maintenance, optimizes production workflows, and enhances overall operational efficiency **[12]**.

Figure 1: The pillars of industry 4.0

2.2 Big Data and Analytics

Big Data analytics involves systematic processing and examination of vast amounts of data to uncover patterns, correlations, and trends that may not be apparent through traditional data analysis methods. Within Industry 4.0, Big Data technologies enable manufacturers to optimize production processes, enhance quality control, and implement predictive maintenance by analyzing data from multiple sources such as sensors, machines, and supply chain systems. By leveraging advanced analytical techniques, companies can make informed, real-time decisions that lead to increased efficiency, reduced operational costs, and improved product innovation **[13]**.

2.3 Cloud Computing

Cloud computing provides scalable and flexible solutions for data storage, management, and processing by utilizing remote servers accessible over the Internet. In Industry 4.0, cloud-based platforms enable seamless integration of data across different systems and locations, facilitating real-time collaboration and remote management of industrial operations. This connectivity is essential for supporting advanced manufacturing systems, such as smart factories, where data-driven processes depend on continuous, reliable access to information. Cloud computing also enhances cybersecurity measures and data backup, ensuring operational continuity **[14]**.

2.4 Artificial Intelligence (AI) and Machine Learning (ML)

Artificial Intelligence (AI) refers to the capability of computer systems to perform tasks that typically require human intelligence, such as speech recognition, decision-making, problem-solving, and natural language understanding. AI encompasses various subfields, including Machine Learning, computer vision, and natural language processing, and relies on simulating human cognitive processes through algorithms and computational models. Machine Learning (ML)**,** a subset of AI, focuses on the design and development of algorithms that can learn from data. Instead of being explicitly programmed to perform a task, ML systems use historical data to train models and improve their performance over time. Applications of ML include image recognition, fraud detection, and predictive maintenance **[15]**.

2.5 Cyber-Physical Systems (CPS)

Cyber-Physical Systems (CPS) refer to the integration of physical processes with computational models, enabling real-time interaction between machines, networks, and humans. These systems form the backbone of Industry 4.0 by creating a digital representation, or "digital twin," of physical assets. CPS can monitor, analyze, and control industrial operations, allowing for precise, adaptive, and automated management of manufacturing processes. This synergy between the physical and digital worlds enhances productivity, reduces errors, and improves safety **[16]**.

2.6 Augmented Reality (AR) / Virtual Reality (VR)

Augmented Reality (AR) and Virtual Reality (VR) technologies provide immersive experiences that overlay digital information onto the physical environment or create entirely simulated environments. In industrial settings, AR/VR tools are used for training, maintenance, product design, and remote assistance. They enable workers to visualize complex data, interact with machinery in a virtual setting, and practice troubleshooting scenarios without disrupting actual production lines. These technologies support safer, more efficient, and faster learning curves, particularly for complex assembly tasks or equipment maintenance **[17]**.

3 AI-Driven Data Management in Mining Supply Chain

3.1 Mining supply chain

The mining supply chain encompasses a series of interconnected stages that guide raw materials from extraction to their final delivery as processed products. Each step plays a crucial role in ensuring efficient operations, optimizing resource management, and meeting market demands. The process (Figure 2) begins with extraction, where raw materials are mined from the earth through methods such as surface mining, underground mining, or placer mining. This stage involves removing the desired ore from the surrounding rock or soil, which can vary significantly depending on the type of mineral deposit and its geographical location **[18]**. Following extraction, the next stage is loading, where the mined materials are collected and transported from the mining site to processing facilities. This step requires efficient handling to ensure that materials are quickly and safely moved to the next phase of the supply chain. The processing stage involves crushing, grinding, and various chemical treatments to extract valuable minerals from the raw material. This is a critical phase in the supply chain as it determines the quality and purity of the final product. Processing can include techniques such as flotation, leaching, and smelting, which are tailored to the specific characteristics of the ore being processed **[19].**

After processing, the refined materials are placed in storage facilities. Storage serves as a buffer to manage supply and demand, allowing companies to hold inventory until it is ready to be transported. Proper storage practices are necessary to ensure the quality of the processed materials and to prepare for the final stages of the supply chain **[20].**

The transport stage involves the movement of refined materials from storage facilities to distribution points or customers. This can include transportation by road, rail, sea, or a combination of these modes. Efficient transport logistics are crucial for ensuring timely delivery and minimizing costs associated with moving heavy and bulky materials **[21].**

Finally, the last step in the mining supply chain is delivery, where the processed and transported materials are delivered to end-users, manufacturers, or further processing plants. This phase is essential for fulfilling contracts, meeting customer demand, and ensuring the overall success of the mining operation **[22].**

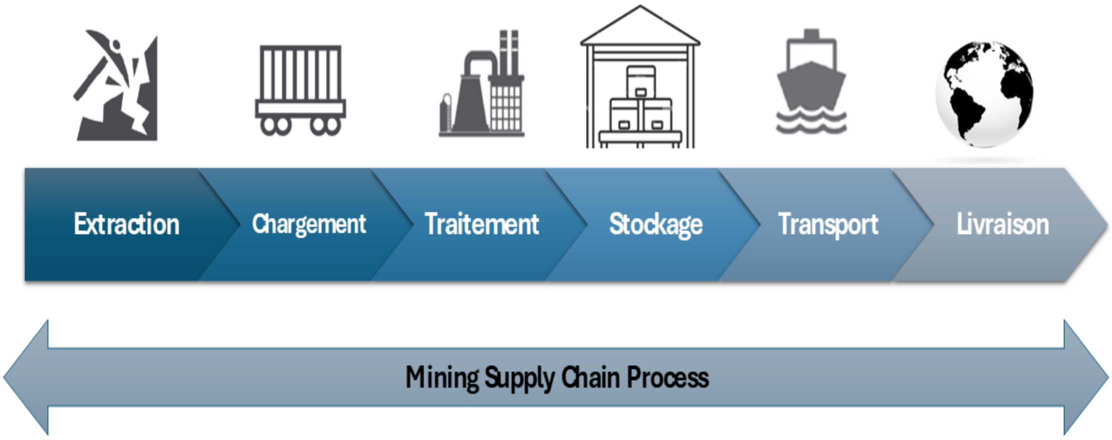


Figure 2: Mining Supply Chain

3.2 Mining supply chain optimization

Mining supply chain optimization involves enhancing the efficiency and effectiveness of the processes that transport resources from extraction to market. This optimization is critical given the inherent complexity of mining operations, which involve multiple stakeholders, diverse geographies, and fluctuating market conditions. Furthermore, the growing emphasis on sustainable practices has added another layer of complexity, requiring innovative solutions to balance economic viability with environmental and social responsibilities. As a result, researchers and industry practitioners have developed various methodologies, including mathematical programming, simulation models, and decentralized approaches, to address these multifaceted challenges. Mathematical programming techniques have proven particularly effective. For instance, Mixed Integer Linear Programming (MILP) has been employed to optimize large-scale mining supply chains, achieving measurable results such as a 4.47% improvement in throughput at a major copper mine **[23]**. This approach allows for precise allocation of resources and optimization of operational schedules, making it an invaluable tool for improving efficiency. Additionally, a bi-level programming model has been developed to handle the complexities of decentralized supply chains. By integrating stakeholder interactions, this model enhances decision-making processes and achieves superior performance using hybrid algorithms **[24].** Simulation models also play a vital role in addressing sustainability challenges. System dynamics-based simulations assess supply chain performance while minimizing economic, social, and environmental costs. This capability is particularly important in industries like copper mining, where sustainability is increasingly prioritized **[25]**. By providing insights into long-term impacts and potential trade-offs, simulation models support strategic planning and informed decision-making.

Technological advancements further enhance supply chain optimization. Grid computing and data mining algorithms streamline the construction of efficient supply chains by optimizing resource allocation and reducing dependency on suppliers. These technologies enable real-time data processing and predictive analytics, providing a competitive edge in managing complex operations **[26]**; **[27]**.

While these methodologies and technologies offer promising results, achieving a balance between efficiency and sustainability remains a significant challenge. Ongoing research and collaboration among stakeholders are essential to overcome these obstacles and ensure that mining supply chains can meet the demands of a rapidly evolving industry.

3.3 AI and Advanced Data Analytics for Mining Supply Chain Optimization

The integration of AI and data analytics in mining supply chains has led to significant improvements across various stages, from extraction to final delivery. One of the key areas is **predictive maintenance**, where AI systems analyze real-time data from equipment sensors to predict potential failures. By identifying issues before they lead to breakdowns, companies can reduce unplanned downtime, extend equipment lifespan, and optimize maintenance schedules. For example, machine learning models can detect abnormal patterns in vibration or temperature data, signaling the need for preventive maintenance **[28]**. This not only ensures the continuous operation of critical machinery but also reduces costs associated with emergency repairs.

Another major application lies in logistics optimization. Efficient supply chain logistics are vital for the mining sector, where transportation costs can be a significant part of the overall expenses. AI-based models can analyze variables such as demand forecasts, weather conditions, and traffic patterns to optimize routes, reduce fuel consumption, and minimize transit times. Additionally, these models can assist in strategic planning by predicting supply and demand fluctuations, enabling better inventory management and stock control **[29]**. This leads to reduced waste and enhanced efficiency throughout the supply chain. AI also plays a crucial role in ore extraction and processing. Machine learning algorithms can process vast amounts of geological data to identify ore bodies and determine optimal extraction methods. By analyzing drill and core sample data, AI can recommend specific techniques (such as flotation or leaching) to maximize recovery rates and minimize energy consumption **[30]**. Moreover, AI systems can dynamically adjust mining operations in response to changing site conditions, ensuring that resources are extracted as efficiently and sustainably as possible. Furthermore, smart production flow management leverages data analytics to optimize operations in processing plants. By monitoring and adjusting production processes in real-time, companies can achieve greater control over product quality, reduce waste, and enhance throughput. Advanced systems equipped with AI can predict potential bottlenecks in production lines and automatically adjust parameters to maintain smooth operation. This kind of real-time decision-making reduces delays and maximizes the use of available resources **[31]**. In addition to these practical applications, strategic resource planning has also benefited from AI. Companies are using advanced algorithms to simulate various production scenarios and forecast long-term resource availability. This allows mining companies to make informed decisions about resource allocation, capital investment, and workforce management. By evaluating multiple scenarios, businesses can identify the most cost-effective and sustainable strategies to meet production targets **[27].** AI-driven insights help companies mitigate risks, adapt to market changes, and align their operations with environmental regulations, contributing to overall sustainability.

4 Enhancing Safety Protocols in Mining through AI, Real-Time Monitoring, and Data Visualization

The integration of AI, real-time monitoring, and data visualization has become a pivotal component in enhancing safety protocols within the mining industry. These advanced technologies enable mining operations to proactively identify potential hazards, mitigate risks, and ensure the well-being of workers in high-risk environments. With the inherent dangers associated with mining, such as exposure to toxic gases, temperature fluctuations, and structural instability, continuous and accurate monitoring is essential for maintaining a safe working environment **[32].**

4.1 Real-Time Monitoring Systems

The deployment of Internet of Things (IoT) devices has revolutionized how mining operations monitor critical safety parameters. IoT sensors are placed throughout mining sites to continuously track vital conditions, such as gas concentrations, temperature, humidity, and equipment status. By leveraging real-time data, these systems can detect abnormal conditions that may indicate the presence of hazardous gases like methane or carbon monoxide, allowing for immediate intervention **[33, 34].** Additionally, IoT networks can monitor the structural integrity of tunnels and equipment, ensuring that any signs of stress or malfunction are addressed promptly. Furthermore, automated alert systems enhance the efficiency of safety protocols. When dangerous conditions are detected, automated systems can instantly send alerts to miners and management teams, ensuring that timely action is taken to prevent accidents. These alerts can be configured to trigger responses such as equipment shutdowns, evacuation procedures, or the activation of emergency ventilation systems, minimizing the risk of injury **[35].** This proactive approach to safety management is critical in environments where even a few seconds of delayed response can have severe consequences.

4.2 AI and Data Visualization

The integration of AI-driven analysis has further improved safety monitoring by enhancing the precision of hazard detection. AI technologies can analyze patterns in sensor data to identify trends that might indicate potential safety issues, such as increasing concentrations of dust or toxic gases. Unlike traditional systems, which rely on threshold-based alerts, AI can predict hazardous conditions based on historical data, leading to more accurate assessments and enabling preventive measures before a crisis develops **[36, 37].** Machine learning models can also optimize safety protocols by learning from past incidents, continuously improving their ability to identify risk factors.

Data visualization tools play a crucial role in helping decision-makers interpret complex datasets. These tools convert real-time sensor data into clear and intuitive visual formats, such as graphs, heat maps, and dashboards. By providing a comprehensive view of safety conditions, data visualization aids in understanding trends and anomalies, allowing managers to make informed decisions on the spot **[38].** For instance, the executive dashboard (see Figure 3). displayed in the image highlights various key performance indicators (KPIs) related to mining operations, such as critical key performance indicators (KPIs) for mining operations, such as production levels, energy, gas, and oil outputs, facilitating quick insights into operational performance and safety standards. Effective visualizations also make it easier to communicate critical information to teams, ensuring that everyone is aware of the safety conditions and necessary precautions.

5 Conclusion

The integration of AI-driven data management and analytics in the mining sector represents a transformative leap forward, driven by the principles of Industry 4.0. These advanced technologies are reshaping the mining landscape by enhancing safety protocols, improving operational efficiency, promoting sustainability, and enabling more precise resource management. Through real-time monitoring and predictive maintenance, Industry 4.0 tools ensure that safety risks are minimized and machinery operates efficiently, reducing downtime and operational costs. The use of AI in optimizing processes such as ore recovery and energy consumption fosters more sustainable practices, helping mining companies address environmental concerns and comply with increasingly stringent regulations.

Moreover, Industry 4.0 facilitates better workforce development through immersive training technologies like AR and VR, equipping employees with the necessary skills to manage and operate advanced systems safely and effectively. These advancements contribute to the overall competitiveness of mining companies by streamlining operations, reducing costs, and improving sustainability, making them more appealing to investors and stakeholders in a global market.

However, the adoption of these technologies does not come without challenges. One of the main obstacles is the high initial investment costs associated with setting up Industry 4.0 infrastructure, which can be prohibitive, especially for smaller mining companies. Additionally, the skills gap is a significant barrier; many current employees may lack the expertise required to operate and maintain these advanced systems, and finding skilled personnel can be challenging.

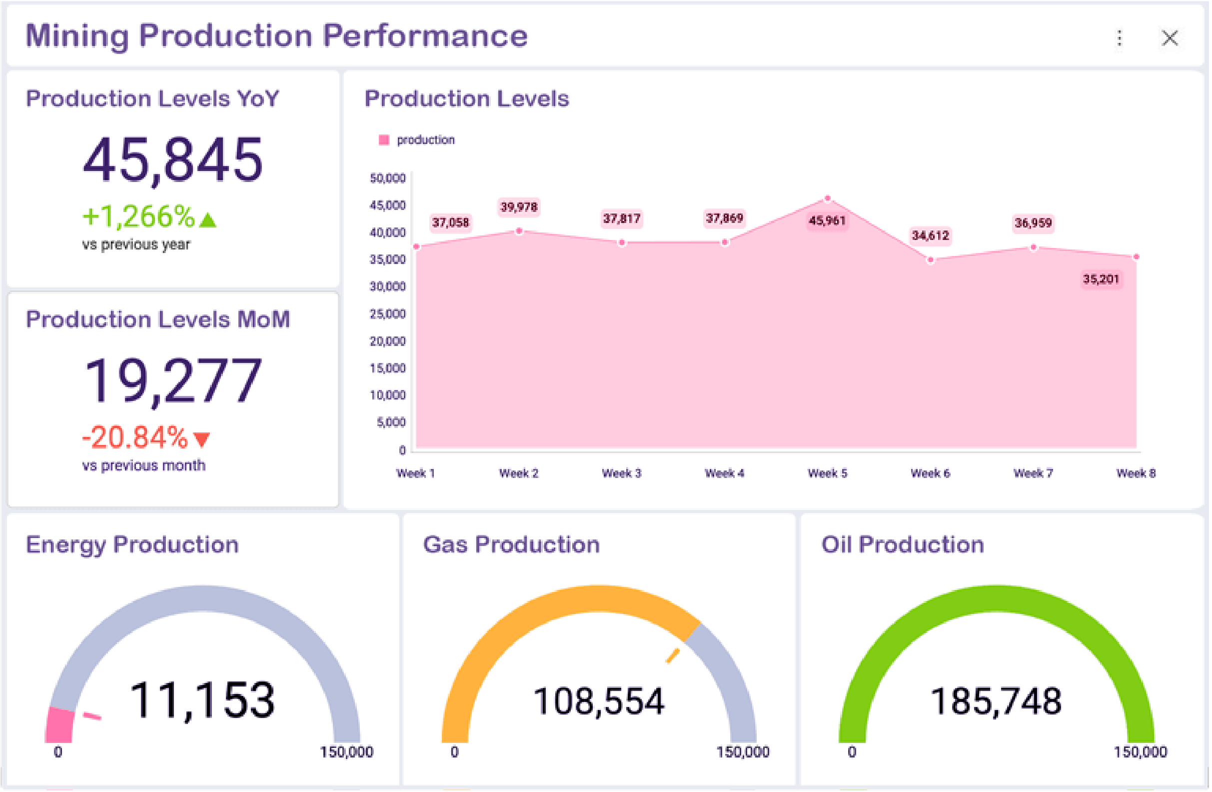


Figure 3: Example of mining industry dashboard

Another concern is data privacy and security, as the vast amounts of sensitive data collected through IoT devices and other digital systems need to be protected from potential breaches. Regulatory hurdles also pose a challenge, as existing laws may not fully accommodate the integration of new technologies, making it difficult for companies to implement changes swiftly. Furthermore, integration with legacy systems can slow down the adoption process, as older equipment may not be compatible with newer Industry 4.0 technologies, requiring costly upgrades or replacements.

Lastly, managing the overwhelming volume of data generated by these systems is a critical challenge. Effective data analytics and management strategies are essential to extract meaningful insights and avoid being overwhelmed by the sheer quantity of information. Risk management, including cyber security and system reliability, also becomes increasingly important as more aspects of mining operations rely on interconnected digital systems.

In conclusion, while challenges exist, the benefits of Industry 4.0 in mining are substantial. Companies that can effectively integrate these technologies will be better positioned to ensure safer, more efficient, and sustainable operations. The future of mining lies in embracing digital transformation, and the ongoing collaboration between the industry and technology developers will play a crucial role in driving this evolution forward.

References

[1] Zulkiflee Abdul-Samad, Lim Liang Xin, Wesam Salah Alaloul, Hafez Salleh. Towards Industrial Revolution (IR) 4.0 in the construction industry: Readiness of contractors. *Results in Engineering*, Volume 22, 2024.

[2] Marta Sukiennik. Challenges Faced by Businesses in the Mining Industry in the Context of the Industry 4.0 Philosophy, *Sciendo*, MAPE 2018, volume 1, issue 1, pp. 621-626, 2018.

[3] Khan, S. A., Alshahrani, S. M., & Khan, S. Z. *Role of digital transformation in the mining industry: A review of the literature. Journal of Cleaner Production*, 263, 121469, 2020.

[4] Luthra, S., & Mangla, S. K. Barriers to Industry 4.0 adoption in manufacturing: A systematic literature review. *Journal of Manufacturing Technology Management*, 29(5), 1130-1155, 2018.

[5] Kumar, A., Singh, R., & Kumar, S. Data analytics in mining: A review. *Mining, Metallurgy & Exploration*, 37(4), 751-762, 2020.

[6] Zhong, R. Y., Xu, C., Klotz, E., & Newman, S. T. Intelligent manufacturing in the context of Industry 4.0: A review. *Engineering*, 3(5), 616-630, 2017.

[7] Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, P. How virtualization helps to create a sustainable factory of the future. *International Journal of Industrial Engineering and Management*, 5(2), 73-82, 2014.

[8] Marr, B. Data-Driven Business Models: The New Industry Standard. *Strategic Direction*, 34(7), 1-3, 2018.

[9] Raj, A., Singh, R., & Srivastava, A. Industry 4.0: A review of the literature and future research directions. *International Journal of Productivity and Performance Management*, 69(4), 775-800, 2020.

[10] Fraser, J., Murphy, R., & O'Kane, T. The role of technology in sustainable mining. *Resources Policy*, 58, 151-158, 2018.

[11] Choudhury, S., Pati, U. C., & Singh, D. Adoption of Industry 4.0 technologies in manufacturing: A framework for small and medium enterprises. *Journal of Manufacturing Technology Management*, 32(6), 1082-1105, 2021.

[12] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645-1660, 2013.

[13] Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. "How 'big data' can make a big impact: Findings from a systematic review and a longitudinal case study." *International Journal of Production Economics, 165,* 234-246,2015.

[14] Xu, X. From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing, 28(1),* 75-86, 2012.

[15] Jordan, M. I., & Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260, 2015.

[16] Monostori, L. "Cyber-physical production systems: Roots, expectations and R&D challenges." *Procedia CIRP, 17,* 9-13, 2014.

[17] Flatt, H., Schreiber, S., Kramer, M., & Lukowicz, P. "AR as a platform for AI-based industrial quality control." *Procedia CIRP, 91,* 9-14, 2020.

[18] Hartman, H. L., & Mutmansky, J. M. Introductory Mining Engineering. John Wiley & Sons, 2002.

[19] Wills, B. A., & Finch, J. A. Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery. 8th Edition. Butterworth-Heinemann, 2015.

[20] Ripinsky, S. Storage and Inventory Management in the Mining Industry. *McGraw-Hill Education,* 2018.

[21] Dale, R. Logistics and Transportation in the Mining Sector. Springer, 2012.

[22] Torries T. F. Economics of the Mineral Industry: From Mining to Market. *Society for Mining, Metallurgy & Exploration*, 1998.

[23] Linh, Mai., Zenn, Saw. Optimization of large mining supply chains through mathematical programming, 165-182, 2024.

[24] Qiang, Zhang., Shi, Qiang, Liu., Andrea, D'Ariano., Sai-Ho, Chung., Mahmoud, Masoud., Xiangong, Li. A bi-level programming methodology for decentralized mining supply chain network design. Expert Systems With Applications, 2024.

[25] Pablo, Becerra., J., Mula., Raquel, Sanchis., Francisco, Campuzano-Bolarín. Simulation optimisation of a sustainable copper mining closed-loop supply chain. International Journal of Systems Science, 2024.

[26] Mei, Xie. Supply Chain Construction and Optimization Model Based on Grid Computing and Process Data Mining Algorithms, 2022.

[27] Vilma, Brodnik. Supply Chain Construction and Optimization Model Based on Grid Computing and Process Data Mining Algorithms, 2022.

[28] Smith, J., & Lee, K. Predictive Maintenance Using Machine Learning for Mining Equipment. *Journal of Mining Science and Technology, 34(5), 425-437,* 2021.

[29] Gupta, R., & Martinez, P. AI-Driven Logistics Optimization in the Mining Supply Chain. *International Journal of Supply Chain Management*, 9(2), 102-115, 2020.

[30] Pereira, L., & Zhang, Y. Smart Mining: Machine Learning Applications for Efficient Ore Extraction. *Mining Engineering Journal*, 57(8), 211-223, 2022.

[31] Allen, C., & Ng, H. Data Analytics for Optimizing Production Flows in Mining Operations. *Journal of Industrial Engineering and Automation*, 12(4), 355-367, 2023.

[32] Brown, D., & Watson, R. (2021). Strategic Resource Planning in the Mining Industry Using Artificial Intelligence. *Resources Policy*, 70, 101894.

[33] Dasari, H. V., Reddy, L. R., Kumar, C. A., & Satyanarayana, A. N. (2024). Advanced safety monitoring system for coal mines using IoT. *International Journal of Innovative Science and Research Technology*, May 4, 2024.

[34] Sumit Kumar Srivastava, Real Time Monitoring System for Mine Safety using Wireless Sensor Network (Multi-Gas Detector), Department Of Mining Engineering National Institute Of Technology Rourkela-769 008, 2015.

[35] Vasanthamma, H., Rathore, L. S., Hussaini, S. S. P., Inayatulla, & Vijay, T. Mining safety and health monitoring system. *International Journal of Advanced Research in Computer and Communication Engineering, 13*(4), 2024.

[36] Arucapalli, Ananth, Chaitanya, P.Rahul., Ananthula, Siddhartha., Mohan, Dholvan., Article, Info. IoT-Based Mining Safety and Air QuBlynk System. Journal of Telecommunication, Electronic and Computer Engineering, 2023.

[37] Daria, Trubicina., Kirill, Varnavskiy., Alexander, Ermakov., Fedor, Nepsha., Roman, Kozlov., Naser, Golsanami., Sergey, Zhironkin. Application of Methods of Artificial Intelligence in Systems for

Continuous Automatic Monitoring of Dust Concentration and Deposits in Mine Atmosphere. arXiv.org, abs/2401.16823, 2024.

[38] Wenhui, Zhang. (2024). Research on the Application of Intelligent Technology in Mining Safety Monitoring. Highlights in Science Engineering and Technology, 106:565-569, 2024.

An analytical review on supply chain risk management : tools for different phases of the process

Nabil KAYOUH[1], Btissam DKHISSI [2]
[1] National school of applied sciences in Tétouan,
Abdelmalek Essaadi university,
nabil.kayouh@etu.uae.ac.ma
[2] National school of applied sciences in Tétouan,
Abdelmalek Essaadi university,
b.dkhissi@uae.ac.ma

Keywords : *supply chain risk management, optimization, quantitative tools , processes*

1 Introduction

Over the past two decades, several global events have disrupted supply chains. Globalization has increased the risk of supply chain disruptions, heightened outsourcing, and emphasized efficiency and lean management [1][2]. The COVID-19 pandemic, with its border closures, lockdowns, and workforce limitations, caused major supply chain disruptions, leading to logistical challenges and reduced production capacity [3]. Other examples include 9/11 [4], Somali pirate attacks, the Fukushima earthquake [5], demand fluctuations, lead time variability, and exchange rate volatility [6]. Extreme weather events, economic forces like the US-China trade war and Brexit, and the Russia-Ukraine conflict have further intensified supply chain challenges [8].Supply chain risk management (SCRM) involves identifying, assessing, mitigating, and monitoring risks at both macro and micro levels through collaborative efforts using qualitative and quantitative techniques [3][9]. It has become an interdisciplinary field with contributions from various domains such as management and engineering [10]. Over the last two decades, research on supply chain risk has significantly evolved [7], leaving a gap for further exploration in both research and practice [2].This paper reviews qualitative-quantitative methodologies in supply chain risk management. Using the PRISMA method, we systematically review the literature to establish a typology, framework, and research agenda for risk factors, mitigation strategies, and tools used in SCRM. We outline the methodology, analyze publication trends by country and field, and evaluate the methods proposed for each stage, concluding with future research directions.

2 Materials and methods

The PRISMA Statement, developed by 29 reviewers, provides a 27-item checklist and a four-phase flow diagram [11]. In this study, we applied PRISMA for the systematic review, identifying, screening, and filtering articles based on research themes, removing duplicates, and retaining relevant ones [12].We started by defining keywords: TITLE-ABS-KEY ("supply chain risk management" OR "SCRM") AND ("quantitative" OR "mathematical model" OR "optimization" OR "simulation" OR "decision"). Only English-language journal articles from January 1, 2012, to November 20, 2022, were considered, yielding 746 articles after applying eligibility criteria. For eligibility, we reviewed titles and abstracts in an Excel file from the Scopus database, marking articles as ‘pass’ or ‘no pass.’ Articles with doubts were fully reviewed. Both qualitative and quantitative studies were included, focusing on risks and SCRM strategies. A total of 316 articles passed and were entered into the Zotero reference management system for further analysis.

3 Results

In this phase, we present the main results of the PRISMA method. First, we analyze publications by country/continent to identify regions most engaged in this topic. Then, we provide a detailed analysis of the quantitative and quali-quantitative models used in each step of SCRM.

3.1 Publication by country

The data extraction has provided insight into the countries with the highest volume of studies in the field of supply chain risk management. India and the United States lead the way with over 51 publications each, making up 16.19% of the total output for both nations. China follows as the third-highest contributor with 43 publications, accounting for 13.65% of the total articles produced over the past decade. Iran has also made significant contributions to this theme, with more than 18 publications. The majority of conducted studies are concentrated in Asia, which stands as the primary contributor to global research output in this field. Following Asia, the United States (North America) and Europe emerge, with both continents having some ground to cover to match the research output of Asian countries, particularly since they comprise developed and industrialized nations. In contrast, Africa and South America significantly trail behind other continents, mainly due to the prevalence of developing and third-world countries within them. These regions would benefit from a greater emphasis on research and increased investment in studying their supply chains (see figure 1).

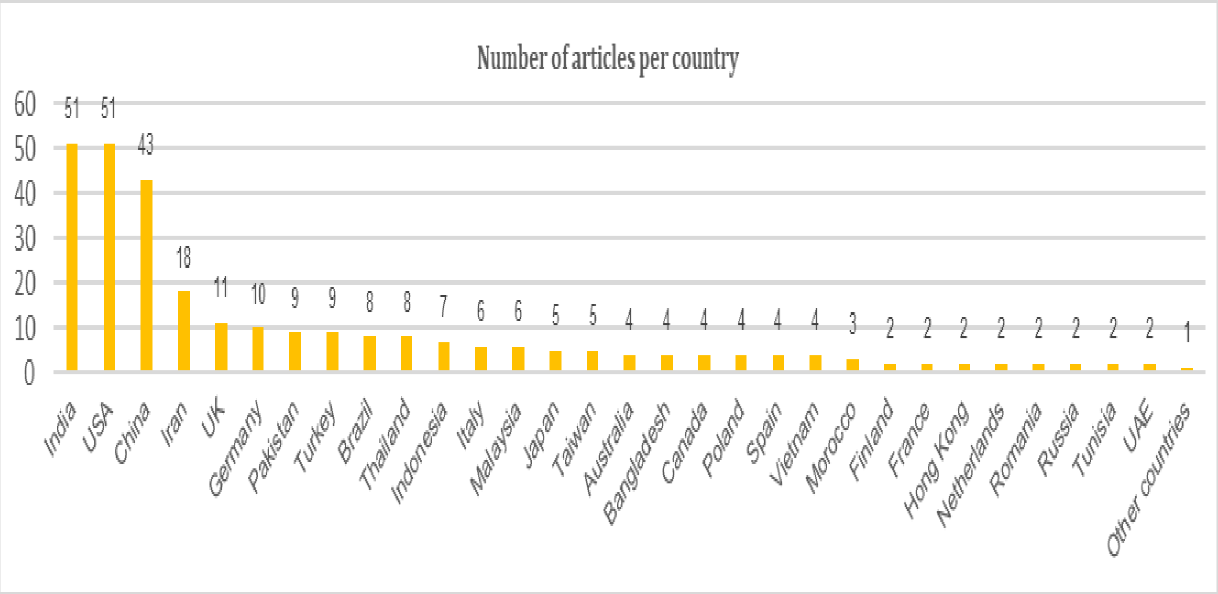


FIG.1 Number of articles per country

3.2 Tools for supply chain risk management

Our analysis identified 235 articles using quantitative or quali-quantitative models in supply chain risk management. The 'SCR identification' phase has 10 articles (3.9%), 'SCRM analysis' 13 articles (5%), 'SCRM assessment or evaluation' 104 articles (41%), and 'SCRM strategy mitigation' 124 articles (49%). The 'SCRM monitoring' phase is underrepresented, with only 1 article (0.3%). The assessment and mitigation phases dominate, covering 90% of the studies, while the identification, analysis, and monitoring phases receive little attention. Future research should focus on these underexplored stages. Multi-criteria decision-making methods are commonly used across the phases, alongside optimization, simulation, and stochastic models. Exploring methodologies specific to one phase and applying them to others is also recommended.

3.2.1 Tools for supply chain risk identification

This phase focuses on identifying risk factors in supply chains, with studies targeting sectors like automotive, agribusiness, and manufacturing. For instance, [13] used hierarchical holographic modeling, while [14] explored IoT methodologies. [15] applied the risk breakdown structure and matrix. Multi-criteria decision-making methods were used in three studies: [16] employed fuzzy ANP and Topsis, and [17][18] used grey-based DEMATEL. Qualitative approaches include HAZOP [19], Interpretive Structural Modeling (ISM) with MICMAC analysis [20], and a qualitative method by [21]. A statistical analysis model was applied by [22]. 55% of articles used multi-criteria decision-making or qualitative methods, with other methodologies used less frequently.

3.2.2 Tools for supply chain risk analysis phase

The analysis phase, which examines the risks identified earlier, employs both multi-criteria decision-making and qualitative models. For example, [23] used a stochastic model with the microscopic Markov chain approach (MMCA) and a two-layer network, while [24] applied business intelligence tools. [25] and [26] used interpretive structural modeling (ISM), and [27] utilized text-mining algorithms. [15] applied the risk breakdown structure and matrix, [28] developed a risk assessment index system, and [29] used fuzzy and grey DEMATEL. Specific optimization models include mean-variance preferences [30], simulation-based optimization [31], Bayesian networks [32], systemic probability sampling [33], and system dynamics with Monte Carlo simulations [34].Our study indicates a consistent use of various methodologies, with optimization models, multi-criteria decision-making approaches, and qualitative models appearing in two articles each. This phase requires further investigation and a broader range of methodological approaches, as robust risk analysis enhances assessment relevance and leads to better strategies. The development of more qualitative-quantitative methods is recommended.

3.2.3 Tools for supply chain risk assessment phase

The supply chain risk assessment phase is crucial in SCRM, with many articles employing various methodologies. Half of these methodologies utilize multi-criteria decision-making (MCDM) methods, primarily the AHP methodology developed by Saaty, which appears in twelve articles [35]. Variants of AHP include fuzzy AHP [36], Pythagorean fuzzy AHP [37], orders-of-magnitude AHP [38][39], neutrosophic AHP [40], and AHP noisy-OR [41]. TOPSIS and DEMATEL follow AHP in usage, with TOPSIS referenced by [42][43] and fuzzy TOPSIS by [44][45]. DEMATEL is highlighted in [17], while fuzzy DEMATEL is used by [46] and [47]. The best-worst method, introduced by Rezaei in 2015, is cited four times [48][49], followed by the VIKOR method [50][51] and its fuzzy variant [48]. Other notable methods include the analytical network process [52][16] and ELECTRE TRI [53][54]. Various MCDM methods are mentioned only once, such as Shannon fuzzy entropy [55], information entropy weight [56], and PROMETHEE [62]. Both fuzzy logic methods and probabilistic graphical models are mentioned ten times, with the Bayesian Belief Network being the most common [71][72].

Statistical models also feature, with authors using different computational approaches [73][74][75][76]. Qualitative methods are used six times, including the house of risk model [77] and MICMAC [78]. Mathematical models appear five times, with various authors utilizing distinct methodologies [80][70][81][82][83]. Additionally, heuristic methods and algorithms are proposed [84][85][86][87], and optimization models are cited four times [88][89][90][67]. Other methodologies are mentioned once, such as the House of Quality method [91], hierarchical holographic modeling [66], and dynamic fault trees [96].

Stochastic models were mentioned by [99] who used a Stochastic integer linear programming approach and [96] who used a Markov chain model and Monte Carlo simulation. Finally [100] made a new proposition of an improved FMEA. In the past decade, the risk assessment phase has been extensively studied, employing various quantitative and quali-quantitative tools. Many articles have used methodological combinations, such as dual MCDM methods or integrating MCDM with mathematical models. This study highlights a strong reliance on MCDM methods due to their importance in assessing risks. We recommend further developing these methodologies for evaluating logistical risks and exploring the integration of established methods with emerging machine learning and AI techniques.

3.2.4 Tools for risk mitigation based on strategies :

After assessing risks in the supply chain, the next step is to propose strategies for risk mitigation. Extensive research over the past decade aims to identify effective combinations of these strategies across various sectors. Multi-criteria decision-making (MCDM) methodologies play a significant role in this area. Specifically, the TOPSIS method is the most widely used approach, as highlighted in six studies [101][59]. The fuzzy version of TOPSIS has also been employed extensively [102][103]. Following TOPSIS, the AHP method has been used by [104][105], with its fuzzy variant applied by [106]. The analytical network process has been utilized by [107][108], while its fuzzy version has been adopted by [109][16]. The DEMATEL method was implemented by [110][17], along with its fuzzy variant [111]. The Best-Worst method was applied in significant studies by [112][49]. Other MCDM methods were used less frequently, including CRITIC [59], FMEA [113], Fuzzy Delphi [103], ELECTRE [114], grey relational analysis [115], Grey layered theory [108], neutrosophic theory [116], Fuzzy MULTIMOORA [106], SWARA [60], and others such as the Three-stage supply chain game decision model [117] and Random multi-attribute decision-making [118].

This phase also included various optimization models, with fifteen proposed overall. Some notable models include those presented by [121], who proposed their own models, [122] with a non-linear Mean-Variance robust optimization problem, and [123], who focused on simulation-based optimization using experimental designs and statistical analyses. Additionally, [124] used Bow-Tie analysis and optimization techniques, while [125] created a robust optimization (RO) approach. Budget-constrained optimization was addressed by [126], who developed a model based on cost and time optimization. Mathematical models ranked third, with [128] proposing risk mitigation models, [110] utilizing grey theory, [129] offering detailed mathematical formulations, [130] employing suitable optimization models, [131] applying a Game-theoretical approach, and [132] suggesting mixed integer mathematical programming models.

Some studies have also considered qualitative approaches. For example, [133] utilized qualitative methodologies, [134] employed complex systems analysis, and [135] took a multiple case-based approach. Additionally, [136] focused on conceptual theory building, [137] proposed Diagraph-matrix methodologies, [138] applied information processing theory, and [139] based his study on real options theory. Stochastic programming models are commonly used, including a Stochastic chance-constrained programming model [140], a stochastic integer linear programming approach [99], and a multi-objective mixed-integer stochastic programming model [141]. Other notable works include a Stochastic optimization model [123], multi-stage stochastic programs [142][143], and a combination of distributionally robust and traditional REI approaches [144]. Stochastic fuzzy multi-objective programming [145], two-stage stochastic programming [146], and the SIMANP metaheuristic algorithm [114] have also been proposed, along with a multi-objective stochastic model [147].

Statistical models were prevalent in this phase of supply chain risk management, with all studies employing structural equation modeling [148]. Some utilized nonparametric statistical methods [149], while others proposed basic data analysis based on descriptive statistics and analysis of variance [150]. Conditional value-at-risk (CVaR) methodologies were introduced [151], along with survey-based methodologies [152], designs of experiments (DOE) [102], and systematic literature analyses (SLA) [153]. Many simulation models were also proposed; for instance, [154] and [155] conducted data and simulation analyses, and [139] performed a computer simulation study. Algorithms were applied, with [156] proposing analytical models, [87] utilizing data-processing rules, and [157] adopting decision tree approaches. Decision support systems were introduced, with [63] proposing a decision-making framework, [158] using decision-making under uncertainty (EUT), [159] developing a decision support system through correspondence analysis and fuzzy inference systems, and [160] creating a rational decision-making technique. Similar to the evaluation phase, fuzzy theory has been adopted by several authors, including [161][162], who used interval-valued fuzzy numbers, and [163], who applied a fuzzy-order function. Programming models were employed four times in this phase, with studies based on mixed integer non-linear programming [164], linear programming [165], mixed integer programming formulations [125], and multi-stage mixed integer programming models [166]. Finally, the latest methodologies, including artificial intelligence, were explored by [167][168].

Several researchers have proposed multi-objective optimization methods to address issues in this phase. For example, [169] utilized a multi-objective optimization model based on the NSGA-III algorithm, while [170] employed a fuzzy multi-objective decision-making approach using the epsilon constraint method. Other approaches include cost-benefit analysis for Pareto-optimal solutions [87], a qualitative approach for multi-objective optimization [112], and a multi-objective optimization-based simulation [171]. Multi-criteria optimization was proposed by [172], who applied linear constrained convex multi-criteria optimization and relevant algorithms.Heuristic models were featured in two studies: [173] proposed a heuristic algorithm based on a relaxation method for decision variables, and [127] introduced another heuristic method. Machine learning algorithms were explored by [174][168]. Multi-criteria decision analysis (MCDA) methods were employed in two studies, with [158] using swing weights and [61] applying a general MCDA method. Non-parametric methods like Data Envelopment Analysis (DEA) were utilized by [104][149], and [175] proposed network non-parametric DEA in his study.Other methodologies were mentioned less frequently, such as the Information Processing Theory (IPT) used by [176], a multi-method quantitative study by [177], and scale purification and factor analysis by [178]. Additionally, simulation for risk mitigation strategies was proposed by [179], and a system dynamics approach was used by [180]. Decision support systems (DSS) were introduced by [181], while [15] discussed the Hierarchical chart model, risk breakdown structure, and risk breakdown matrix. Other models included integrated multi-criteria decision-making approaches [182], IoT frameworks [126], linear programming models [183], Morphological Analysis (MA) [184], PLS-SEM modeling [185], optimized neural networks [90], resource-based view (RBV) models [186], innovative portfolio approaches [146], probabilistic graphical models [187], and fuzzy cognitive mapping [188]. Additionally, quality function deployment (QFD) was proposed by [182], the structural analysis method MICMAC was utilized [189], and expected value and chance-constrained models were developed using uncertain interval programming techniques [190].

It is evident that multi-criteria decision-making methods are the most frequently used in this phase, accounting for 25% of cases, followed closely by optimization models and various mathematical and stochastic models. This trend highlights the critical importance of selecting appropriate models at this stage of supply chain risk management to ensure effective risk mitigation strategies.

The selection of risk mitigation strategies has been a significant focus in studies over the past decade, aiming to identify proactive and reactive measures to minimize damage and mitigate risks. The COVID-19 crisis has exposed vulnerabilities in global supply chains, emphasizing the need for innovative solutions to reduce impacts. Utilizing quantitative models to prioritize these methods is essential. We recommend focusing on the latest methodologies in multi-criteria decision-making, mathematical models, and optimization techniques, as well as leveraging advanced technologies in artificial intelligence, blockchain, and Industry 4.0 to explore more effective solutions for this phase.

3.2.5 Tools for supply chain monitoring phase

The final phase of the Supply Chain Risk Management (SCRM) process is risk monitoring. In this context, [191] proposed a study on supply chain risk governance, integrating risk governance within the Risk Management Support System (RMSS) domain and collaborative SCRM into a multilevel framework. This framework encompasses the supply network, risk management procedures, and inter-organizational governance mechanisms, aiming to enhance collaboration in SCRM. This study is unique in its focus on the monitoring phase, highlighting the need for further research to investigate this phase and develop new models that ensure the sustainability of results achieved through mitigation strategies.

4 Conclusion and perspectives

This paper presents a systematic review of the literature on supply chain risk management (SCRM) from 2012 to 2023. The review was prompted by several factors, notably the significant impact of the COVID-19 crisis and the Russia-Ukraine war on supply chains, highlighting the need for updated analysis. Additionally, the study aims to explore risk response strategies and existing quantitative tools in the literature from a fresh perspective.

Using the PRISMA methodology and focusing on a selective set of journals, our analysis reviewed 316 articles on SCRM published during the specified period, addressing four key research questions: which countries have contributed most to supply chain risk management, and what tools are employed at each stage of the SCRM process? Proper identification of risk factors allows for the development of effective strategies to minimize or anticipate potential damage.

The findings indicate that the initial and final phases of the SCRM process are the least represented in the literature, with only 5% of articles focusing on risk analysis. In contrast, the assessment and mitigation phases garnered significant attention, comprising 90% of the articles, which predominantly proposed quantitative or qualitative-quantitative models.

A common aspect across all five phases of SCRM is the incorporation of at least one qualitative approach, with multi-criteria decision-making (MCDM) techniques utilized in almost every phase. Optimization, simulation, and stochastic models are also considered in the analysis, assessment, and mitigation stages.

Most studies are concentrated in Asia, which accounts for a significant portion of global research, followed by the United States and Europe. The latter regions must intensify efforts to catch up with their Asian counterparts, especially given their developed and industrialized status.

Based on the literature review, several future research directions are proposed. Firstly, while global quantitative tools are essential, there is a need to expand research in less developed countries. Most studies have been conducted in Asia and Europe, and regions such as Africa, Australia, and the Americas should enhance their contributions. Secondly, certain SCRM stages, particularly risk identification, analysis, and monitoring, require more attention, as our study revealed a lack of quantitative models for these phases.

In addition to the commonly used methodologies like MCDM and mathematical optimization models, further investigation into new combinations of methodologies is recommended. Lastly, the literature has underrepresented small companies, necessitating a greater focus on their challenges in future research.

Despite our efforts to justify each analytical step, limitations exist in this study. Primarily, our focus was restricted to the Scopus database, which, while extensive, does not encompass all historical publications.

References

[1] K. E. Stecke et S. Kumar, « Sources of Supply Chain Disruptions, Factors That Breed Vulnerability, and Mitigating Strategies », *J. Mark. Channels*, vol. 16, n° 3, p. 193-226, juin 2009, doi: 10.1080/10466690902932551.

[2] M. V. C. Fagundes, E. O. Teles, S. A. B. Vieira De Melo, et F. G. M. Freires, « Supply chain risk management modelling: A systematic literature network analysis review », *IMA J. Manag. Math.*, vol. 31, n° 4, p. 387-416, sept. 2020, doi: 10.1093/imaman/dpaa019.

[3] A. Wicaksana, W. Ho, S. Talluri, et A. Dolgui, « A decade of progress in supply chain risk management: risk typology, emerging topics, and research collaborators », *Int. J. Prod. Res.*, vol. 60, n° 24, p. 7155-7177, déc. 2022, doi: 10.1080/00207543.2022.2077672.

[4] K. J. Mizgier, S. M. Wagner, et M. P. Jüttner, « Disentangling diversification in supply chain networks », *Int. J. Prod. Econ.*, vol. 162, p. 115-124, avr. 2015, doi: 10.1016/j.ijpe.2015.01.007.

[5] G. Olcott et N. Oliver, « Social Capital, Sensemaking, and Recovery: Japanese Companies and the 2011 Earthquake », *Calif. Manage. Rev.*, vol. 56, n° 2, p. 5-22, févr. 2014, doi: 10.1525/cmr.2014.56.2.5.

[6] M. Esmaeilikia, B. Fahimnia, J. Sarkis, K. Govindan, A. Kumar, et J. Mo, « A tactical supply chain planning model with multiple flexibility options: an empirical evaluation », *Ann. Oper. Res.*, vol. 244, n° 2, p. 429-454, sept. 2016, doi: 10.1007/s10479-013-1513-2.

[7] M. Pournader, A. Kach, et S. (Sri) Talluri, « A Review of the Existing and Emerging Topics in the Supply Chain Risk Management Literature », *Decis. Sci.*, vol. 51, n° 4, p. 867-919, août 2020, doi: 10.1111/deci.12470.

[8] « Russia-Ukraine war impact on supply chains and inflation ». Consulté le: 21 octobre 2023. [En ligne]. Disponible sur: https://kpmg.com/us/en/articles/2022/russia-ukraine-war-impact-supply-chains-inflation.html

[9] Y. H. Smith Jeremy, « In U.S.-China Trade War, New Supply Chains Rattle Markets », Carnegie Endowment for International Peace. Consulté le: 21 mai 2023. [En ligne]. Disponible sur: https://carnegieendowment.org/2020/06/24/in-u.s.-china-trade-war-new-supply-chains-rattle-markets-pub-82145

[10] N. S. Hudin, N. F. Habidin, P. Surya, et K. C. Hussin, « Mitigation Strategies in Supply Chain Risk Management: A Literature Review », vol. 22, n° 5, 2021.

[11] A. Liberati *et al.*, « The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration », *PLoS Med.*, vol. 6, n° 7, p. e1000100, juill. 2009, doi: 10.1371/journal.pmed.1000100.

[12] S. Anwar, T. Perdana, M. Rachmadi, et T. I. Noor, « Traceability Information Model for Sustainability of Black Soybean Supply Chain: A Systematic Literature Review », *Sustainability*, vol. 14, n° 15, p. 9498, août 2022, doi: 10.3390/su14159498.

[13] J. Hou et X. Zhao, « Toward a supply chain risk identification and filtering framework using systems theory », *Asia Pac. J. Mark. Logist.*, vol. 33, n° 6, p. 1482-1497, juin 2021, doi: 10.1108/APJML-05-2020-0342.

[14] Q. Gao, S. Guo, X. Liu, G. Manogaran, N. Chilamkurti, et S. Kadry, « Simulation analysis of supply chain risk management system based on IoT information platform », *Enterp. Inf. Syst.*, vol. 14, n° 9-10, p. 1354-1378, nov. 2020, doi: 10.1080/17517575.2019.1644671.

[15] A. C. Cagliano, A. De Marco, S. Grimaldi, et C. Rafele, « An integrated approach to supply chain risk analysis », *J. Risk Res.*, vol. 15, n° 7, p. 817-840, août 2012, doi: 10.1080/13669877.2012.666757.

[16] S. Aghajanian et R. I. Shevchenko-Perepy, « An Empirical Procurement Risk Management Framework in Supply Chain Networks : A Hybrid Approach », *Ind. Eng. Manag. Syst.*, vol. 17, n° 4, p. 730-744, déc. 2018, doi: 10.7232/iems.2018.17.4.730.

[17] S. Mithun Ali, Md. A. Moktadir, G. Kabir, J. Chakma, Md. J. U. Rumi, et Md. T. Islam, « Framework for evaluating risks in food supply chain: Implications in food wastage reduction », *J. Clean. Prod.*, vol. 228, p. 786-800, août 2019, doi: 10.1016/j.jclepro.2019.04.322.

[18] A. Prashar et S. Aggarwal, « Modeling enablers of supply chain quality risk management: a grey-DEMATEL approach », *TQM J.*, vol. 32, n^o 5, p. 1059-1076, oct. 2019, doi: 10.1108/TQM-05-2019-0132.

[19] Mitkowski P.T. et Zenka-Podlaszewska D., « Hazop method in identification of risks in a cpfr supply chain », *Chem. Eng. Trans.*, vol. 39, p. 445-450, août 2014, doi: 10.3303/CET1439075.

[20] S. SerdarAsan et S. Karadayı-Usta, « DEFINING AND MODELING RISKS IN SERVICE SUPPLY CHAINS », 국제산업공학저널, vol. 28, n^o 1, oct. 2021, doi: 10.23055/IJIETAP.2021.28.1.5605.

[21] F. R. Azmi, A. Abdullah, E. R. Cahyadi, H. Musa, et J. R. Sa'ari, « Type of Risk in Halal Food Supply Chain: A Review », vol. 9, n^o 4, 2020.

[22] H. Rogers, M. Srivastava, K. S. Pawar, et J. Shah, « Supply chain risk management in India – practical insights », *Int. J. Logist. Res. Appl.*, vol. 19, n^o 4, p. 278-299, juill. 2016, doi: 10.1080/13675567.2015.1075476.

[23] L. Huo, H. Guo, Y. Cheng, et X. Xie, « A new model for supply chain risk propagation considering herd mentality and risk preference under warning information on multiplex networks », *Phys. Stat. Mech. Its Appl.*, vol. 545, p. 123506, mai 2020, doi: 10.1016/j.physa.2019.123506.

[24] P.-J. Wu et P. Chaipiyaphan, « Diagnosis of delivery vulnerability in a logistics system for logistics risk management », *Int. J. Logist. Manag.*, vol. 31, n^o 1, p. 43-58, déc. 2019, doi: 10.1108/IJLM-02-2019-0069.

[25] S. M. Zanjirchi, N. Jalilian, et A. Mirhoseini, « Risk-agility interactive model: a new look at agility drivers », *J. Model. Manag.*, vol. 12, n^o 4, p. 690-711, nov. 2017, doi: 10.1108/JM2-01-2016-0007.

[26] D.-W. Kwak, V. S. Rodrigues, R. Mason, S. Pettit, et A. Beresford, « Risk interaction identification in international supply chain logistics: Developing a holistic model », *Int. J. Oper. Prod. Manag.*, vol. 38, n^o 2, p. 372-389, févr. 2018, doi: 10.1108/IJOPM-03-2016-0121.

[27] M. R. Sheikhattar, N. Nezafati, et S. Shokouhyar, « A thematic analysis–based model for identifying the impacts of natural crises on a supply chain for service integrity: a text analysis approach », *Environ. Sci. Pollut. Res.*, vol. 29, n^o 52, p. 79413-79433, nov. 2022, doi: 10.1007/s11356-022-21380-x.

[28] M. Junaid, Y. Xue, et M. W. Syed, « Construction of Index System for Risk Assessment in Supply Chains of Automotive Industry », vol. 9, n^o 4, 2020.

[29] K.-J. Wu, C.-J. Liao, M.-L. Tseng, M. K. Lim, J. Hu, et K. Tan, « Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties », *J. Clean. Prod.*, vol. 142, p. 663-676, janv. 2017, doi: 10.1016/j.jclepro.2016.04.040.

[30] S. Mukherjee et S. S. Padhi, « Sourcing decision under interconnected risks: an application of mean–variance preferences approach », *Ann. Oper. Res.*, vol. 313, n^o 2, p. 1243-1268, juin 2022, doi: 10.1007/s10479-021-04485-3.

[31] D. Bandaly, A. Satir, et L. Shanker, « Impact of lead time variability in supply chain risk management », *Int. J. Prod. Econ.*, vol. 180, p. 88-100, oct. 2016, doi: 10.1016/j.ijpe.2016.07.014.

[32] S. Cao, K. Bryceson, et D. Hine, « An Ontology-based Bayesian network modelling for supply chain risk propagation », *Ind. Manag. Data Syst.*, vol. 119, n^o 8, p. 1691-1711, sept. 2019, doi: 10.1108/IMDS-01-2019-0032.

[33] M. S. Shahbaz, R. Z. RM Rasi, et M. F. Bin Ahmad, « A novel classification of supply chain risks: Scale development and validation », *J. Ind. Eng. Manag.*, vol. 12, n^o 1, p. 201, avr. 2019, doi: 10.3926/jiem.2792.

[34] Hiroki Okubo, Shigeji Miyazaki, Takahiro Ogura, et Hisahi Onari, « Scenario-based simulation for supply chain risk management », *International information institute*, p. 4705-4716, 2013.

[35] Q. Meng, X. Liu, Y. Song, et W. Wang, « An extended generalized TODIM method for risk assessment of supply chain in social commerce under interval type-2 fuzzy environment », *J. Intell. Fuzzy Syst.*, vol. 37, n^o 6, p. 8551-8565, déc. 2019, doi: 10.3233/JIFS-190061.

[36] K. Ganguly et G. Kumar, « Supply Chain Risk Assessment: a fuzzy AHP approach », *Oper. Supply Chain Manag. Int. J.*, p. 1-13, févr. 2019, doi: 10.31387/oscm0360217.

[37] S. Lahane et R. Kant, « Evaluation and ranking of solutions to mitigate circular supply chain risks », *Sustain. Prod. Consum.*, vol. 27, p. 753-773, juill. 2021, doi: 10.1016/j.spc.2021.01.034.

[38] J. Silva, C. Araujo, et L. Marques, « Siloed Perceptions in Pharmaceutical Supply Chain Risk Management: A Brazilian Perspective », *Lat. Am. Bus. Rev.*, vol. 21, nº 3, p. 223-254, juill. 2020, doi: 10.1080/10978526.2020.1731315.

[39] P. Raghuram, P. Sandeep, V. R. Sreedharan, et T. Saikouk, « Development of a supply chain risk mitigation index for distillery », *TQM J.*, vol. 33, nº 3, p. 618-639, avr. 2021, doi: 10.1108/TQM-01-2020-0008.

[40] M. Abdel-Basset, M. Gunasekaran, M. Mohamed, et N. Chilamkurti, « RETRACTED: A framework for risk assessment, management and evaluation: Economic tool for quantifying risks in supply chain », *Future Gener. Comput. Syst.*, vol. 90, p. 489-502, janv. 2019, doi: 10.1016/j.future.2018.08.035.

[41] A. F. H. Librantz, I. Costa, M. de M. Spinola, G. C. de Oliveira Neto, et L. Zerbinatti, « Risk assessment in software supply chains using the Bayesian method », *Int. J. Prod. Res.*, vol. 59, nº 22, p. 6758-6775, nov. 2021, doi: 10.1080/00207543.2020.1825860.

[42] Wang Y. et Hao H.X., « Research on the supply chain risk assessment of the fresh agricultural products based on the improved toptsis algorithm », *Chem. Eng. Trans.*, vol. 51, p. 445-450, juill. 2016, doi: 10.3303/CET1651075.

[43] N. Bohlouli, A. Shahbazpour, et M. R. Gorbanian Tabrizi, « Identify and Prioritize Supply Chain Risks by TOPSIS (Case study: Motorsazan Company) », *Australian Journal of Basic and Applied Sciences*, p. 406-412, 2012.

[44] C. Sun, Y. Xiang, S. Jiang, et Q. Che, « A supply chain risk evaluation method based on fuzzy topsis », *Int. J. Saf. Secur. Eng.*, vol. 5, nº 2, p. 150-161, juin 2015, doi: 10.2495/SAFE-V5-N2-150-161.

[45] A. Samvedi, V. Jain, et F. T. S. Chan, « Quantifying risks in a supply chain through integration of fuzzy AHP and fuzzy TOPSIS », *Int. J. Prod. Res.*, vol. 51, nº 8, p. 2433-2442, avr. 2013, doi: 10.1080/00207543.2012.741330.

[46] A. Üstündağ, S. Çıkmak, M. Çankaya Eyiol, et M. C. Ungan, « Evaluation of supply chain risks by fuzzy DEMATEL method: a case study of iron and steel industry in Turkey », *Int. J. Prod. Manag. Eng.*, vol. 10, nº 2, p. 195-209, juill. 2022, doi: 10.4995/ijpme.2022.17169.

[47] A. Samvedi et V. Jain, « A study on the interactions between supply chain risk management criteria using fuzzy DEMATEL method », *Int. J. Oper. Res.*, vol. 18, nº 3, p. 255, 2013, doi: 10.1504/IJOR.2013.056685.

[48] M. Bayanati, A. Peivandizadeh, M. R. Heidari, S. Foroutan Mofrad, M. R. Sasouli, et A. Pourghader Chobar, « Prioritize Strategies to Address the Sustainable Supply Chain Innovation Using Multicriteria Decision-Making Methods », *Complexity*, vol. 2022, p. 1-11, oct. 2022, doi: 10.1155/2022/1501470.

[49] P. Pakdeenarong et T. Hengsadeekul, « Supply chain risk management of organic rice in Thailand », *Uncertain Supply Chain Manag.*, p. 165-174, 2020, doi: 10.5267/j.uscm.2019.7.007.

[50] M. Nazam, L. Yao, M. Hashim, S. A. Baig, et M. K. Khan, « The Application of a Multi-Attribute Group Decision Making Model Based on Linguistic Extended VIKOR for Quantifying Risks in a Supply Chain Under a Fuzzy Environment »:, *Int. J. Inf. Syst. Supply Chain Manag.*, vol. 13, nº 2, p. 27-46, avr. 2020, doi: 10.4018/IJISSCM.2020040102.

[51] Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand, D. Sumrit, et S. Srisawad, « Fuzzy failure mode and effect analysis model for operational supply chain risks assessment: an application in canned tuna manufacturer in Thailand », *Logforum*, vol. 18, nº 1, p. 77-96, mars 2022, doi: 10.17270/J.LOG.2022.645.

[52] P. K. Tarei, J. J. Thakkar, et B. Nag, « A hybrid approach for quantifying supply chain risk and prioritizing the risk drivers: A case of Indian petroleum supply chain », *J. Manuf. Technol. Manag.*, vol. 29, nº 3, p. 533-569, mars 2018, doi: 10.1108/JMTM-10-2017-0218.

[53] K. Govindan et M. B. Jepsen, « Supplier risk assessment based on trapezoidal intuitionistic fuzzy numbers and ELECTRE TRI-C: a case illustration involving service suppliers », *J. Oper. Res. Soc.*, vol. 67, nº 2, p. 339-376, févr. 2016, doi: 10.1057/jors.2015.51.

[54] S. Carpitella, I. Mzougui, et J. Izquierdo, « Multi-criteria risk classification to enhance complex supply networks performance », *OPSEARCH*, vol. 59, nº 3, p. 769-785, sept. 2022, doi: 10.1007/s12597-021-00568-8.

[55] A. Shahin, A. Kianersi, et A. Shali, « Prioritizing Key Supply Chain Risks Using the Risk Assessment Matrix and Shannon Fuzzy Entropy — with a Case Study in the Home Appliance Industry », *J. Adv. Manuf. Syst.*, vol. 17, nº 03, p. 333-351, sept. 2018, doi: 10.1142/S0219686718500208.

[56] C. Liu, H. Ji, et J. Wei, « Smart Supply Chain Risk Assessment in Intelligent Manufacturing », *J. Comput. Inf. Syst.*, vol. 62, nº 3, p. 609-621, mai 2022, doi: 10.1080/08874417.2021.1872045.

[57] S. Prakash, G. Soni, et A. P. S. Rathore, « A grey based approach for assessment of risk associated with facility location in global supply chain », *Grey Syst. Theory Appl.*, vol. 5, nº 3, p. 419-436, nov. 2015, doi: 10.1108/GS-12-2014-0059.

[58] S. Althaf et C. W. Babbitt, « Disruption risks to material supply chains in the electronics sector », *Resour. Conserv. Recycl.*, vol. 167, p. 105248, avr. 2021, doi: 10.1016/j.resconrec.2020.105248.

[59] R. Rostamzadeh, M. K. Ghorabaee, K. Govindan, A. Esmaeili, et H. B. K. Nobar, « Evaluation of sustainable supply chain risk management using an integrated fuzzy TOPSIS-CRITIC approach », *J. Clean. Prod.*, vol. 175, p. 651-669, févr. 2018, doi: 10.1016/j.jclepro.2017.12.071.

[60] M. Yazdani, E. D. R. S. Gonzalez, et P. Chatterjee, « A multi-criteria decision-making framework for agriculture supply chain risk management under a circular economy context », *Manag. Decis.*, vol. 59, nº 8, p. 1801-1826, août 2021, doi: 10.1108/MD-10-2018-1088.

[61] H. Mulyati et J. Geldermann, « Managing risks in the Indonesian seaweed supply chain », *Clean Technol. Environ. Policy*, vol. 19, nº 1, p. 175-189, janv. 2017, doi: 10.1007/s10098-016-1219-7.

[62] S. P. Venkatesan et S. Kumanan, « Supply chain risk prioritisation using a hybrid AHP and PROMETHEE approach », *Int. J. Serv. Oper. Manag.*, vol. 13, nº 1, p. 19, 2012, doi: 10.1504/IJSOM.2012.048274.

[63] M. Er Kara, S. Ü. Oktay Fırat, et A. Ghadge, « A data mining-based framework for supply chain risk management », *Comput. Ind. Eng.*, vol. 139, p. 105570, janv. 2020, doi: 10.1016/j.cie.2018.12.017.

[64] L. Wang, « A multi-level fuzzy comprehensive assessment for supply chain risks », *J. Intell. Fuzzy Syst.*, vol. 41, nº 4, p. 4947-4954, nov. 2021, doi: 10.3233/JIFS-189981.

[65] F. Aqlan et S. S. Lam, « A fuzzy-based integrated framework for supply chain risk assessment », *Int. J. Prod. Econ.*, vol. 161, p. 54-63, mars 2015, doi: 10.1016/j.ijpe.2014.11.013.

[66] D. Nakandala, H. Lau, et L. Zhao, « Development of a hybrid fresh food supply chain risk assessment model », *Int. J. Prod. Res.*, vol. 55, nº 14, p. 4180-4195, juill. 2017, doi: 10.1080/00207543.2016.1267413.

[67] F. Aqlan, « A software application for rapid risk assessment in integrated supply chains », *Expert Syst. Appl.*, vol. 43, p. 109-116, janv. 2016, doi: 10.1016/j.eswa.2015.08.028.

[68] A. Ahadi Oroumieh, « Supply chain risk assessment of the Iranian mining industry by using fuzzy inference system », *Uncertain Supply Chain Manag.*, vol. 3, nº 3, p. 273-282, 2015, doi: 10.5267/j.uscm.2015.3.003.

[69] Z. Fang et Q. Wang, « Cross-Border E-Commerce Supply Chain Risk Evaluation with FUZZY-ISM Model », *Secur. Commun. Netw.*, vol. 2021, p. 1-14, sept. 2021, doi: 10.1155/2021/5155524.

[70] S. Zhou, C. Hu, Y. Xie, et W. Chang, « Research on supply chain risk assessment with intuitionistic fuzzy information », *J. Intell. Fuzzy Syst.*, vol. 30, nº 6, p. 3367-3372, avr. 2016, doi: 10.3233/IFS-152084.

[71] A. Qazi et M. C. E. Simsekler, « Worst Expected Best method for assessment of probabilistic network expected value at risk: application in supply chain risk management », *Int. J. Qual. Reliab. Manag.*, vol. 39, nº 1, p. 155-175, janv. 2022, doi: 10.1108/IJQRM-07-2020-0238.

[72] S. Anwar, T. Djatna, Sukardi, et P. Suryadarma, « Modelling supply chain risks and their impacts on the performance of the sago starch agro-industry », *Int. J. Product. Perform. Manag.*, vol. 71, nº 6, p. 2361-2392, juin 2022, doi: 10.1108/IJPPM-10-2020-0556.

[73] K. J. Mizgier, « Global sensitivity analysis and aggregation of risk in multi-product supply chain networks », *Int. J. Prod. Res.*, vol. 55, nº 1, p. 130-144, janv. 2017, doi: 10.1080/00207543.2016.1198504.

[74] D. Liang, R. Bhamra, Z. Liu, et Y. Pan, « Risk Propagation and Supply Chain Health Control Based on the SIR Epidemic Model », *Mathematics*, vol. 10, nº 16, p. 3008, août 2022, doi: 10.3390/math10163008.

[75] S. Dechprom et K. Jermsittiparsert, « Sustainability Related Supply Chain Risks: A Case of Multiple Organizational Strategic Networks », *Int. J. Innov.*, vol. 5, nº 2, 2019.

[76] H. Q. Truong et (Y. H., « Supply chain risk management: manufacturing- and service-oriented firms », *Journal of Manufacturing Technology Management*, p. 218-239, 2018.

[77] W. Boonyanusith et P. Jittamai, « Blood Supply Chain Risk Management using House of Risk Model », *Walailak J. Sci. Technol. WJST*, vol. 16, nº 8, p. 573-591, janv. 2018, doi: 10.48048/wjst.2019.3472.

[78] B. Harish, B. Prabhas, et A. Anil Kumar, « Analysis and model development of supply chain risk variables in the Indian manufacturing context) », *International Journal of Business Continuity and Risk Management*, International Journal of Business Continuity and Risk Management, p. 307-329, 2020.

[79] L. Luo, G. Qiping Shen, G. Xu, Y. Liu, et Y. Wang, « Stakeholder-Associated Supply Chain Risks and Their Interactions in a Prefabricated Building Project in Hong Kong », *J. Manag. Eng.*, vol. 35, nº 2, p. 05018015, mars 2019, doi: 10.1061/(ASCE)ME.1943-5479.0000675.

[80] Pei, Xue-liang, Tian, Ye-zhuang, et Sun, Hong-yi, « A risk assessment study on enterprise supply chain management based on extension matter element », *Advances in Information Sciences and Service Sciences*, p. 31-39.

[81] J. Nakatani *et al.*, « A graph theory-based methodology for vulnerability assessment of supply chains using the life cycle inventory database », *Omega*, vol. 75, p. 165-181, mars 2018, doi: 10.1016/j.omega.2017.03.003.

[82] R. Rajesh et V. Ravi, « Analyzing drivers of risks in electronic supply chains: a grey–DEMATEL approach », *Int. J. Adv. Manuf. Technol.*, vol. 92, nº 1-4, p. 1127-1145, sept. 2017, doi: 10.1007/s00170-017-0118-3.

[83] X. Chen, J. Luo, X. Wang, et D. Yang, « Supply chain risk management considering put options and service level constraints », *Comput. Ind. Eng.*, vol. 140, p. 106228, févr. 2020, doi: 10.1016/j.cie.2019.106228.

[84] R. Handfield, H. Sun, et L. Rothenberg, « Assessing supply chain risk for apparel production in low cost countries using newsfeed analysis », *Supply Chain Manag. Int. J.*, vol. 25, nº 6, p. 803-821, juin 2020, doi: 10.1108/SCM-11-2019-0423.

[85] H. Jiao, W. Ding, Y. Shi, N. Zhao, et B. Wang, « Supply chain risk assessment model based on cloud model with subjective preference weight allocation algorithm », *Intell. Decis. Technol.*, vol. 14, nº 2, p. 133-142, juill. 2020, doi: 10.3233/IDT-180001.

[86] S. Lu, « Enterprise supply chain risk assessment based on improved neural network algorithm and machine learning », *J. Intell. Fuzzy Syst.*, vol. 40, nº 4, p. 7013-7024, avr. 2021, doi: 10.3233/JIFS-189532.

[87] A. Qazi et P. Akhtar, « Risk matrix driven supply chain risk management: Adapting risk matrix based tools to modelling interdependent risks and risk appetite », *Comput. Ind. Eng.*, vol. 139, p. 105351, janv. 2020, doi: 10.1016/j.cie.2018.08.002.

[88] SUHARJITO et MARIMIN, « RISKS BALANCING MODEL OF AGRI-SUPPLY CHAIN USING FUZZY RISKS UTILITY REGRESSION », *Journal of Theoretical and Applied Information Technology*, p. 134-144, 2012.

[89] Xiaofen Ji et Wei Chen, « Evaluation of Supply Chain Risk and Its Empirical Analysis for Apparel Processing Enterprises », *Journal of Convergence Information Technology*, p. 44-53, 2012.

[90] F. Jianying, Y. Bianyu, L. Xin, T. Dong, et M. Weisong, « Evaluation on risks of sustainable supply chain based on optimized BP neural networks in fresh grape industry », *Comput. Electron. Agric.*, vol. 183, p. 105988, avr. 2021, doi: 10.1016/j.compag.2021.105988.

[91] R. Sukwadi et A. Caesar, « An integrated approach for supply chain risk management », *Eng. Manag. Prod. Serv.*, vol. 14, n° 1, p. 38-48, mars 2022, doi: 10.2478/emj-2022-0004.

[92] A. Alora et M. K. Barua, « An integrated structural modelling and MICMAC analysis for supply chain disruption risk classification and prioritisation in India ».

[93] A. Malik, K. I. A. Khan, S. Qayyum, F. Ullah, et A. Maqsoom, « Resilient Capabilities to Tackle Supply Chain Risks: Managing Integration Complexities in Construction Projects », *Buildings*, vol. 12, n° 9, p. 1322, août 2022, doi: 10.3390/buildings12091322.

[94] A. M. N. Mohib et A. M. Deif, « Supply chain multi-state risk assessment using universal generating function », *Prod. Plan. Control*, vol. 31, n° 9, p. 699-708, juill. 2020, doi: 10.1080/09537287.2019.1680891.

[95] E. Levner et A. Ptuskin, « Entropy-based model for the ripple effect: managing environmental risks in supply chains », *Int. J. Prod. Res.*, vol. 56, n° 7, p. 2539-2551, avr. 2018, doi: 10.1080/00207543.2017.1374575.

[96] X. Lei et C. A. MacKenzie, « Assessing risk in different types of supply chains with a dynamic fault tree », *Comput. Ind. Eng.*, vol. 137, p. 106061, nov. 2019, doi: 10.1016/j.cie.2019.106061.

[97] A. Sibevei, A. Azar, M. Zandieh, S. M. Khalili, et M. Yazdani, « Developing a Risk Reduction Support System for Health System in Iran: A Case Study in Blood Supply Chain Management », *Int. J. Environ. Res. Public. Health*, vol. 19, n° 4, p. 2139, févr. 2022, doi: 10.3390/ijerph19042139.

[98] S. Foli, S. Durst, L. Davies, et S. Temel, « Supply Chain Risk Management in Young and Mature SMEs », *J. Risk Financ. Manag.*, vol. 15, n° 8, p. 328, juill. 2022, doi: 10.3390/jrfm15080328.

[99] G. J. L. Micheli, R. Mogre, et A. Perego, « How to choose mitigation measures for supply chain risks », *Int. J. Prod. Res.*, vol. 52, n° 1, p. 117-129, janv. 2014, doi: 10.1080/00207543.2013.828170.

[100] J. Liu, F. Liu, H. Zhou, et Y. Kong, « An Integrated Method of Supply Chains Vulnerability Assessment », *Sci. Program.*, vol. 2016, p. 1-10, 2016, doi: 10.1155/2016/2819238.

[101] S. Salehi Heidari, M. Khanbabaei, et M. Sabzehparvar, « A model for supply chain risk management in the automotive industry using fuzzy analytic hierarchy process and fuzzy TOPSIS », *Benchmarking Int. J.*, vol. 25, n° 9, p. 3831-3857, nov. 2018, doi: 10.1108/BIJ-11-2016-0167.

[102] M. R. Marjani, M. Habibi, et A. Pazhouhandeh, « An innovative hybrid fuzzy TOPSIS based on design of experiments for multi-criteria supplier evaluation and selection », *International Journal of Operational Research*, p. 171-209, 2022.

[103] X. Wang, P. Tiwari, et X. Chen, « Communicating supply chain risks and mitigation strategies: a comprehensive framework », *Prod. Plan. Control*, vol. 28, n° 13, p. 1023-1036, oct. 2017, doi: 10.1080/09537287.2017.1329562.

[104] H. Min, H. Park, et S. B. Ahn, « Measuring the supply chain risk in offshoring countries using data envelopment analysis and the analytic hierarchy process », *Benchmarking Int. J.*, vol. 24, n° 7, p. 1977-1994, oct. 2017, doi: 10.1108/BIJ-04-2016-0054.

[105] H. Babu, P. Bhardwaj, et A. K. Agrawal, « Assessment and prioritisation of risks and SCRM strategies in Indian small and medium enterprises (SMEs) supply chains': a SCRM-AHP model », *International Journal of Agile Systems and Management*, p. 313-332, 2021.

[106] M. Tavana, A. Shaabani, S. Mansouri Mohammadabadi, et N. Varzgani, « An integrated fuzzy AHP- fuzzy MULTIMOORA model for supply chain risk-benefit assessment and supplier selection », *Int. J. Syst. Sci. Oper. Logist.*, vol. 8, n° 3, p. 238-261, juill. 2021, doi: 10.1080/23302674.2020.1737754.

[107] S. Fazli, R. Kiani Mavi, et M. Vosooghidizaji, « Crude oil supply chain risk management with DEMATEL–ANP », *Oper. Res.*, vol. 15, n° 3, p. 453-480, oct. 2015, doi: 10.1007/s12351-015-0182-0.

[108] R. Rajesh, « A grey-layered ANP based decision support model for analyzing strategies of resilience in electronic supply chains », *Eng. Appl. Artif. Intell.*, vol. 87, p. 103338, janv. 2020, doi: 10.1016/j.engappai.2019.103338.

[109] S. Sinrat et W. Atthirawong, « Integrated factor analysis and fuzzy analytic network process (FANP) model for supplier selection based on supply chain risk factors », *Research Journal of Business Management*, p. 106-123, 2015.

[110] R. Rajesh et V. Ravi, « Modeling enablers of supply chain risk mitigation in electronic supply chains: A Grey–DEMATEL approach », *Comput. Ind. Eng.*, vol. 87, p. 126-139, sept. 2015, doi: 10.1016/j.cie.2015.04.028.

[111] S. Chirra et D. Kumar, « Evaluation of Supply Chain Flexibility in Automobile Industry with Fuzzy DEMATEL Approach », *Glob. J. Flex. Syst. Manag.*, vol. 19, n° 4, p. 305-319, déc. 2018, doi: 10.1007/s40171-018-0195-7.

[112] F. Li, C.-H. Wu, L. Zhou, G. Xu, Y. Liu, et S.-B. Tsai, « A model integrating environmental concerns and supply risks for dynamic sustainable supplier selection and order allocation », *Soft Comput.*, vol. 25, n° 1, p. 535-549, janv. 2021, doi: 10.1007/s00500-020-05165-3.

[113] Nasim Nahavandi et Parizad Tavakoli, « RISK MANAGEMENT OF PROCUREMENT PROCESSES IN AUTOMOTIVE SUPPLY CHAIN; BAHMAN MOTOR COMPANY », *International Journal of Industrial Engineering*, 2022. [En ligne]. Disponible sur: https://doi.org/10.23055/ijietap.2022.29.1.6735

[114] H. Farughi et S. Mostafayi, « A hybrid approach based on ANP, ELECTRE and SIMANP metaheuristic method for outsourcing manufacturing procedures according to supply chain risks - Case study: A medical equipment manufacturer company in Iran », *Decis. Sci. Lett.*, p. 77-94, 2017, doi: 10.5267/j.dsl.2016.6.003.

[115] P. Shojaei et S. A. S. Haeri, « Development of supply chain risk management approaches for construction projects: A grounded theory approach », *Comput. Ind. Eng.*, vol. 128, p. 837-850, févr. 2019, doi: 10.1016/j.cie.2018.11.045.

[116] M. Junaid, Y. Xue, M. W. Syed, J. Z. Li, et M. Ziaullah, « A Neutrosophic AHP and TOPSIS Framework for Supply Chain Risk Assessment in Automotive Industry of Pakistan », *Sustainability*, vol. 12, n° 1, p. 154, déc. 2019, doi: 10.3390/su12010154.

[117] Min Yang, « Supply chain value co-creation and flexibility management decisions taking supply risk into account », *International Journal of Applied Decision Sciences*, p. 221-243, 2022.

[118] S. Shen, « Multi-attribute decision-making methods based on normal random variables in supply chain risk management », *Appl. Math. Nonlinear Sci.*, vol. 7, n° 1, p. 719-728, janv. 2022, doi: 10.2478/amns.2021.2.00147.

[119] Brodetskiy, Gennady L, Gusev, Denis A, et Shidlovskii, Ivan G, « Multi-criteria optimisation under the conditions of uncertainty in logistics and supply chain management », *International Journal of Logistics Systems and Management*, International Journal of Logistics Systems and Management, p. 207-227.

[120] E. Benedito, C. Martínez-Costa, et S. Rubio, « Introducing Risk Considerations into the Supply Chain Network Design », *Processes*, vol. 8, n° 6, p. 743, juin 2020, doi: 10.3390/pr8060743.

[121] Bin Mai et Jianguo Liu, « Information Security Risks Propagation and Management in Supply Chain: An Analytical Approach », *journal of information system security*, p. 3-25, 2019.

[122] P. Kumar Tarei, G. Kumar, et M. Ramkumar, « A Mean-Variance robust model to minimize operational risk and supply chain cost under aleatory uncertainty: A real-life case application in petroleum supply chain », *Comput. Ind. Eng.*, vol. 166, p. 107949, avr. 2022, doi: 10.1016/j.cie.2022.107949.

[123] D. Bandaly, A. Satir, et L. Shanker, « Integrated supply chain risk management via operational methods and financial instruments », *Int. J. Prod. Res.*, vol. 52, n° 7, p. 2007-2025, avr. 2014, doi: 10.1080/00207543.2013.844376.

[124] F. Aqlan et S. S. Lam, « Supply chain risk modelling and mitigation », *Int. J. Prod. Res.*, vol. 53, nº 18, p. 5640-5656, sept. 2015, doi: 10.1080/00207543.2015.1047975.

[125] S. Prakash, S. Kumar, G. Soni, V. Jain, et A. P. S. Rathore, « Closed-loop supply chain network design and modelling under risks and demand uncertainty: an integrated robust optimization approach », *Ann. Oper. Res.*, vol. 290, nº 1-2, p. 837-864, juill. 2020, doi: 10.1007/s10479-018-2902-3.

[126] T. Kieras, J. Farooq, et Q. Zhu, « I-SCRAM: A Framework for IoT Supply Chain Risk Analysis and Mitigation Decisions », *IEEE Access*, vol. 9, p. 29827-29840, 2021, doi: 10.1109/ACCESS.2021.3058338.

[127] C. C. Srinivasan, M. Mohsin Khan, et N. Nirmala S S, « A heuristic method for logistics supply chain risk control of green edible agricultural products geaps in South indian contet », *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, 2019. [En ligne]. Disponible sur: http://www.ijstr.org/final-print/nov2019/A-Heuristic-Method-For-Logistics-Supply-Chain-Risk-Control-Of-Green-Edible-Agricultural-Products-geaps-In-South-indian-Context.pdf

[128] M. D. Sherwin, H. R. Medal, C. A. MacKenzie, et K. J. Brown, « Identifying and mitigating supply chain risks using fault tree optimization », *IISE Trans.*, vol. 52, nº 2, p. 236-254, févr. 2020, doi: 10.1080/24725854.2019.1630865.

[129] M. Goswami, « Supply chain centric product line selection: a functional risk focused approach », *Int. J. Prod. Res.*, vol. 56, nº 20, p. 6678-6700, oct. 2018, doi: 10.1080/00207543.2018.1436788.

[130] F. Dunke, I. Heckmann, S. Nickel, et F. Saldanha-da-Gama, « Time traps in supply chains: Is optimal still good enough? », *Eur. J. Oper. Res.*, vol. 264, nº 3, p. 813-829, févr. 2018, doi: 10.1016/j.ejor.2016.07.016.

[131] K. Papapanagiotou et D. Vlachos, « Managing supply chain disruption risks: a game-theoretical approach », *Int. J. Math. Oper. Res.*, vol. 4, nº 5, p. 515, 2012, doi: 10.1504/IJMOR.2012.048929.

[132] Z. Khojasteh-Ghamari, « Supplier selection in multiple sourcing: a proactive approach to manage the supply chain risk », p. 20.

[133] S. Ambulkar, J. V. Blackhurst, et D. E. Cantor, « Supply chain risk mitigation competency: an individual-level knowledge-based perspective », *Int. J. Prod. Res.*, vol. 54, nº 5, p. 1398-1411, mars 2016, doi: 10.1080/00207543.2015.1070972.

[134] R. C. Basole, M. A. Bellamy, H. Park, et J. Putrevu, « Computational Analysis and Visualization of Global Supply Network Risks », *IEEE Trans. Ind. Inform.*, vol. 12, nº 3, p. 1206-1213, juin 2016, doi: 10.1109/TII.2016.2549268.

[135] V. Dohale, P. Ambilkar, A. Gunasekaran, et P. Verma, « Supply chain risk mitigation strategies during COVID-19: exploratory cases of “make-to-order” handloom saree apparel industries », *Int. J. Phys. Distrib. Logist. Manag.*, vol. 52, nº 2, p. 109-129, mars 2022, doi: 10.1108/IJPDLM-12-2020-0450.

[136] I. Kilubi et H. Rogers, « Bridging the gap between supply chain risk management and strategic technology partnering capabilities: insights from social capital theory », *Supply Chain Manag. Int. J.*, vol. 23, nº 4, p. 278-292, juin 2018, doi: 10.1108/SCM-02-2017-0091.

[137] R. Rajesh, V. Ravi, et R. Venkata Rao, « Selection of risk mitigation strategy in electronic supply chains using grey theory and digraph-matrix approaches », *Int. J. Prod. Res.*, vol. 53, nº 1, p. 238-257, janv. 2015, doi: 10.1080/00207543.2014.948579.

[138] H. Fan, T. C. E. Cheng, G. Li, et P. K. C. Lee, « The Effectiveness of Supply Chain Risk Information Processing Capability: An Information Processing Perspective », *IEEE Trans. Eng. Manag.*, vol. 63, nº 4, p. 414-425, nov. 2016, doi: 10.1109/TEM.2016.2598814.

[139] N. Carbonara et R. Pellegrino, « How do supply chain risk management flexibility-driven strategies perform in mitigating supply disruption risks? ».

[140] L. Liu, S. Li, et Y. Wu, « Supply Chain Risk Management in Chinese Chemical Industry based on Stochastic Chance-Constrained Programming Model », *Appl. Math. Inf. Sci.*, vol. 8, nº 3, p. 1201-1206, mai 2014, doi: 10.12785/amis/080333.

[141] X. Liu, Z. Li, et L. He, « A Multi-objective Stochastic Programming Model for Order Quantity Allocation under Supply Uncertainty », vol. 3, nº 3, 2014.

[142] K. D. Simsek, C. Haksoz, et M. Cakanyildirim, « Managing supply chain risk through take-or-pay gas contracts in the presence of buyers' storage facilities », *J. Energy Mark.*, 2019, doi: 10.21314/JEM.2019.197.

[143] E. Sanci, M. S. Daskin, Y.-C. Hong, S. Roesch, et D. Zhang, « Mitigation strategies against supply disruption risk: a case study at the Ford Motor Company », *Int. J. Prod. Res.*, vol. 60, n° 19, p. 5956-5976, oct. 2022, doi: 10.1080/00207543.2021.1975058.

[144] S. Y. Gao, D. Simchi-Levi, C.-P. Teo, et Z. Yan, « Disruption Risk Mitigation in Supply Chains: The Risk Exposure Index Revisited », *Oper. Res.*, vol. 67, n° 3, p. 831-852, mai 2019, doi: 10.1287/opre.2018.1776.

[145] D. Wu, D. D. Wu, Y. Zhang, et D. L. Olson, « Supply chain outsourcing risk using an integrated stochastic-fuzzy optimization approach », *Inf. Sci.*, vol. 235, p. 242-258, juin 2013, doi: 10.1016/j.ins.2013.02.002.

[146] T. Sawik, « Disruption mitigation and recovery in supply chains using portfolio approach », *Omega*, vol. 84, p. 232-248, avr. 2019, doi: 10.1016/j.omega.2018.05.006.

[147] S. Vahid Nooraie et M. M. Parast, « Mitigating supply chain disruptions through the assessment of trade-offs among risks, costs and investments in capabilities », *Int. J. Prod. Econ.*, vol. 171, p. 8-21, janv. 2016, doi: 10.1016/j.ijpe.2015.10.018.

[148] J. F. Kros, M. Falasca, S. Dellana, et W. J. Rowe, « Mitigating counterfeit risk in the supply chain: an empirical study », *TQM J.*, vol. 32, n° 5, p. 983-1002, déc. 2019, doi: 10.1108/TQM-02-2019-0054.

[149] S. S. Talluri, T. J. Kull, H. Yildiz, et J. Yoon, « Assessing the Efficiency of Risk Mitigation Strategies in Supply Chains », *J. Bus. Logist.*, vol. 34, n° 4, p. 253-269, déc. 2013, doi: 10.1111/jbl.12025.

[150] D. Zimon et P. Madzík, « Standardized management systems and risk management in the supply chain », *Int. J. Qual. Reliab. Manag.*, vol. 37, n° 2, p. 305-327, nov. 2019, doi: 10.1108/IJQRM-04-2019-0121.

[151] C. Fang, X. Liao, et M. Xie, « A hybrid risks-informed approach for the selection of supplier portfolio », *Int. J. Prod. Res.*, vol. 54, n° 7, p. 2019-2034, avr. 2016, doi: 10.1080/00207543.2015.1076947.

[152] M. Madani et K. Jermsittiparsert, « A CONCEPTUAL AND EMPIRICAL MODEL OF SUPPLY CHAIN RISK MANAGEMENT MODEL IN INDONESIAN SMES », *Humanit. Soc. Sci. Rev.*, vol. 7, n° 3, p. 703-710, août 2019, doi: 10.18510/hssr.2019.73101.

[153] P. K. Tarei, J. J. Thakkar, et B. Nag, « Benchmarking the relationship between supply chain risk mitigation strategies and practices: an integrated approach », *Benchmarking Int. J.*, vol. 27, n° 5, p. 1683-1715, avr. 2020, doi: 10.1108/BIJ-12-2019-0523.

[154] G. Hong, « The Risk Elasticity Measure Index Exploration of Supply Chain and Mechanism Research », *Open Cybern. Syst. J.*, vol. 9, n° 1, p. 2497-2502, oct. 2015, doi: 10.2174/1874110X01509012497.

[155] X. Wang, V. S. Rodrigues, et E. Demir, « Managing Your Supply Chain Pantry: Food Waste Mitigation Through Inventory Control », *IEEE Eng. Manag. Rev.*, vol. 47, n° 2, p. 97-102, juin 2019, doi: 10.1109/EMR.2019.2915064.

[156] G. Žigienė, E. Rybakovas, R. Vaitkienė, et V. Gaidelys, « Setting the Grounds for the Transition from Business Analytics to Artificial Intelligence in Solving Supply Chain Risk », *Sustainability*, vol. 14, n° 19, p. 11827, sept. 2022, doi: 10.3390/su141911827.

[157] A. J. Ruiz-Torres, F. Mahmoodi, et A. Z. Zeng, « Supplier selection model with contingency planning for supplier failures », *Comput. Ind. Eng.*, vol. 66, n° 2, p. 374-382, oct. 2013, doi: 10.1016/j.cie.2013.06.021.

[158] A. Qazi, A. Dickson, J. Quigley, et B. Gaudenzi, « Supply chain risk network management: A Bayesian belief network and expected utility based approach for managing supply chain risks », *Int. J. Prod. Econ.*, vol. 196, p. 24-42, févr. 2018, doi: 10.1016/j.ijpe.2017.11.008.

[159] P. K. Tarei, J. J. Thakkar, et B. Nag, « Development of a decision support system for assessing the supply chain risk mitigation strategies: an application in Indian petroleum supply chain », *J. Manuf. Technol. Manag.*, vol. 32, n° 2, p. 506-535, juill. 2020, doi: 10.1108/JMTM-02-2020-0035.

[160] S. Sharma et D. Malhotra, « An integrated rational decision making and inventory management model for managing supply chain risks in EPC: a power sector case », *Int. J. Inf. Decis. Sci.*, vol. 4, n° 2/3, p. 268, 2012, doi: 10.1504/IJIDS.2012.047076.

[161] A. Ghadge, X. Fang, S. Dani, et J. Antony, « Supply chain risk assessment approach for process quality risks », *Int. J. Qual. Reliab. Manag.*, vol. 34, n° 7, p. 940-954, août 2017, doi: 10.1108/IJQRM-01-2015-0010.

[162] K. Chatterjee et S. Kar, « Multi-criteria analysis of supply chain risk management using interval valued fuzzy TOPSIS », *OPSEARCH*, vol. 53, n° 3, p. 474-499, sept. 2016, doi: 10.1007/s12597-015-0241-6.

[163] A. Mostafaeipour, M. Qolipour, et H. Eslami, « Implementing fuzzy rank function model for a new supply chain risk management », *J. Supercomput.*, vol. 73, n° 8, p. 3586-3602, août 2017, doi: 10.1007/s11227-017-1960-7.

[164] R. Ghasemy Yaghin et M. Goh, « Procurement, production, and price planning under visibility and risk: a generalised Benders decomposition method », *Int. J. Prod. Res.*, vol. 59, n° 18, p. 5626-5646, sept. 2021, doi: 10.1080/00207543.2020.1788735.

[165] Massala, Otso A., Neidigh, Robert O., et Langella, Ian M., « Teaching operational flexibility for mitigating risk in global supply chains », *Operations Management Education Review*, p. 87-89, 2020.

[166] Aceng Muhtaram Mirfani *et al.*, « An Integrated Multi-Objective Approach to Managing Supply Risks in a Flexible Supply Chain », *Industrial Engineering & Management Systems*, p. 596-603, 2021.

[167] G. Baryannis, S. Dani, et G. Antoniou, « Predicting supply chain risks using machine learning: The trade-off between performance and interpretability », *Future Gener. Comput. Syst.*, vol. 101, p. 993-1004, déc. 2019, doi: 10.1016/j.future.2019.07.059.

[168] A. Deiva Ganesh et P. Kalpana, « Future of artificial intelligence and its influence on supply chain risk management – A systematic review », *Comput. Ind. Eng.*, vol. 169, p. 108206, juill. 2022, doi: 10.1016/j.cie.2022.108206.

[169] A. Ghadge, S. K. Jena, S. Kamble, D. Misra, et M. K. Tiwari, « Impact of financial risk on supply chains: a manufacturer-supplier relational perspective », *Int. J. Prod. Res.*, vol. 59, n° 23, p. 7090-7105, déc. 2021, doi: 10.1080/00207543.2020.1834638.

[170] M.-C. Yu et M. Goh, « A multi-objective approach to supply chain visibility and risk », *Eur. J. Oper. Res.*, vol. 233, n° 1, p. 125-130, févr. 2014, doi: 10.1016/j.ejor.2013.08.037.

[171] J. Yoon, S. Talluri, H. Yildiz, et W. Ho, « Models for supplier selection and risk mitigation: a holistic approach », *Int. J. Prod. Res.*, vol. 56, n° 10, p. 3636-3661, mai 2018, doi: 10.1080/00207543.2017.1403056.

[172] S.-J. Qu, M. Goh, R. De Souza, et T.-N. Wang, « Proximal Point Algorithms for Convex Multi-criteria Optimization with Applications to Supply Chain Risk Management », *J. Optim. Theory Appl.*, vol. 163, n° 3, p. 949-956, déc. 2014, doi: 10.1007/s10957-014-0540-8.

[173] S. V. Nooraie, « A multi-objective approach to supply chain risk management_ Integrating visibility with supply and demand risk », 2015.

[174] S. Hongjin, « Analysis of risk factors in financial supply chain based on machine learning and IoT technology », *J. Intell. Fuzzy Syst.*, vol. 40, n° 4, p. 6421-6431, avr. 2021, doi: 10.3233/JIFS-189482.

[175] R. Kraude, S. Narayanan, et S. Talluri, « Evaluating the performance of supply chain risk mitigation strategies using network data envelopment analysis », *Eur. J. Oper. Res.*, vol. 303, n° 3, p. 1168-1182, déc. 2022, doi: 10.1016/j.ejor.2022.03.016.

[176] H. S. Birkel et E. Hartmann, « Internet of Things – the future of managing supply chain risks », *Supply Chain Manag. Int. J.*, vol. 25, n° 5, p. 535-548, juin 2020, doi: 10.1108/SCM-09-2019-0356.

[177] A. Saleem, « Automated Inventory Management Systems and its impact on Supply Chain Risk Management in Manufacturing firms of Pakistan », vol. 9, n° 3, 2020.

[178] S. Dellana, W. J. Rowe, et Y. Liao, « A scale for measuring organizational risk management maturity in the supply chain », *Benchmarking Int. J.*, vol. 29, n° 3, p. 905-930, mars 2022, doi: 10.1108/BIJ-11-2020-0578.

[179] J. R. Macdonald, C. W. Zobel, S. A. Melnyk, et S. E. Griffis, « Supply chain risk and resilience: theory building through structured experiments and simulation », *Int. J. Prod. Res.*, vol. 56, nº 12, p. 4337-4355, juin 2018, doi: 10.1080/00207543.2017.1421787.

[180] M. L. Keilhacker et S. Minner, « Supply chain risk management for critical commodities: A system dynamics model for the case of the rare earth elements », *Resour. Conserv. Recycl.*, vol. 125, p. 349-362, oct. 2017, doi: 10.1016/j.resconrec.2017.05.004.

[181] R. Mogre, S. S. Talluri, et F. DAmico, « A Decision Framework to Mitigate Supply Chain Risks: An Application in the Offshore-Wind Industry », *IEEE Trans. Eng. Manag.*, vol. 63, nº 3, p. 316-325, août 2016, doi: 10.1109/TEM.2016.2567539.

[182] C.-H. Hsu, R.-Y. Yu, A.-Y. Chang, W.-L. Liu, et A.-C. Sun, « Applying Integrated QFD-MCDM Approach to Strengthen Supply Chain Agility for Mitigating Sustainable Risks », *Mathematics*, vol. 10, nº 4, p. 552, févr. 2022, doi: 10.3390/math10040552.

[183] R. O. Neidigh et I. M. Langella, « Teaching students supply chain risk using a linear programming model of the fictitious online retailer *Elbe* », *J. Educ. Bus.*, vol. 95, nº 3, p. 148-158, avr. 2020, doi: 10.1080/08832323.2019.1622502.

[184] M. Elmsalmi, W. Hachicha, et A. M. Aljuaid, « Modeling Sustainable Risks Mitigation Strategies Using a Morphological Analysis-Based Approach: A Real Case Study », *Sustainability*, vol. 13, nº 21, p. 12210, nov. 2021, doi: 10.3390/su132112210.

[185] A. Z. Piprani, N. I. Jaafar, S. M. Ali, M. S. Mubarik, et M. Shahbaz, « Multi-dimensional supply chain flexibility and supply chain resilience: the role of supply chain risks exposure », *Oper. Manag. Res.*, vol. 15, nº 1-2, p. 307-325, juin 2022, doi: 10.1007/s12063-021-00232-w.

[186] J. El Baz et S. Ruel, « Can supply chain risk management practices mitigate the disruption impacts on supply chains' resilience and robustness? Evidence from an empirical survey in a COVID-19 outbreak era », *Int. J. Prod. Econ.*, vol. 233, p. 107972, mars 2021, doi: 10.1016/j.ijpe.2020.107972.

[187] M. D. Garvey et S. Carnovale, « The rippled newsvendor: A new inventory framework for modeling supply chain risk severity in the presence of risk propagation », *Int. J. Prod. Econ.*, vol. 228, p. 107752, oct. 2020, doi: 10.1016/j.ijpe.2020.107752.

[188] A. Qazi, J. Quigley, A. Dickson, et Ş. Ö. Ekici, « Exploring dependency based probabilistic supply chain risk measures for prioritising interdependent risks and strategies », *Eur. J. Oper. Res.*, vol. 259, nº 1, p. 189-204, mai 2017, doi: 10.1016/j.ejor.2016.10.023.

[189] M. Elmsalmi, W. Hachicha, et A. M. Aljuaid, « Prioritization of the Best Sustainable Supply Chain Risk Management Practices Using a Structural Analysis Based-Approach », *Sustainability*, vol. 13, nº 9, p. 4608, avr. 2021, doi: 10.3390/su13094608.

[190] Sharmistha Halder Jana, « Application of expected value and chance constraint on uncertain supply chain model with cost, risk and visibility for COVID-19 pandemic », *International Journal of Management Science and Engineering Management*, p. 10-24, 2022.

[191] V. Ahlqvist, A. Norrman, et M. Jahre, « Supply Chain Risk Governance : Towards a Conceptual Multi-Level Framework », *Oper. Supply Chain Manag. Int. J.*, p. 382-395, juill. 2020, doi: 10.31387/oscm0430278.

An NSGA-II algorithm for synchromodal transportation problem considering external costs

Ananthakrishnan Vaikkathe, Abdelhamid Benaini, Jaouad Boukachour
Le Havre University Normandy, France
`{ananthakrishnan.vaikkathe, jaouad.boukachour, abdelhamid.benaini}@univ-lehavre.fr`

Abstract : *This study presents a multi-objective optimization approach for synchromodal transportation, incorporating external costs, a critical factor often overlooked in synchromodal literature. A mathematical model is developed to minimize total duration, carbon emissions, and external costs. The proposed multi-objective optimization eliminates the need for objective normalization, a common requirement in prior studies on synchromodality.*

To solve the problem, a Non-dominated Sorting Genetic Algorithm II (NSGA-II) is employed. The algorithm parameters are calibrated using Taguchi analysis, and the model is validated using instances based on the Seine Axis in France, which features a rich network of inland waterways and terminals. The results demonstrate that NSGA-II achieves optimal solutions for small instances and delivers high-quality solutions with minimal deviation from the best-known solutions for larger instances. In addition, the impact of external costs on mode selection is analyzed, providing valuable insights into sustainable transport planning.

Keywords : *Mixed Integer Linear Programming, Synchromodal Transportation, Multigraph, NSGA-II, Seine Axis.*

1 Introduction

Synchromodal transportation, or synchromodality, represents the latest advancement in multimodal transportation paradigms, evolving from intermodal transportation. It aims to seamlessly integrate various transport modes into a unified system, enabling flexible transitions between them to ensure a coherent, door-to-door logistics solution. In Europe, despite the rise of intermodal transport, uni-modal trucking remains dominant for both short- and long-distance freight, primarily due to its door-to-door flexibility. One of the main challenges faced by intermodal transport is its inherent inflexibility—a limitation that synchromodal transportation is designed to overcome.

Synchromodality uses the concept of amodal booking, providing logistics service providers (LSPs) the flexibility to choose transportation routes and modes based on specific shipment priorities, such as reducing emissions or minimizing transit time. In synchromodality, LSPs generate a set of potential transportation plans, selecting the most suitable one based on a prioritized criteria such as carbon emission, duration, etc. Most existing studies employ a weighted objective function to evaluate transportation routes. However, this study adopts a multi-objective optimization approach to address the synchromodal transportation problem, aiming to generate Pareto-optimal itineraries for efficiently transporting containers between origin and destination.

The transportation modes considered in this study are road, rail, and inland water transport (IWT), and transportation itineraries are evaluated based on three key factors: total duration, total carbon emissions, and total external costs. Although duration and carbon emissions are commonly considered objectives in the synchromodal literature, external costs are often overlooked. External costs refer to the negative impacts of an economic activity on third parties, such as environmental degradation, accidents, or noise pollution. To our knowledge, this is the first study in synchromodality that explicitly incorporates external costs in the selection criteria, thus assessing its impact on mode selection and mode share distribution.

The remainder of the paper is organized as follows. The section 2 reviews the major modeling approaches in the existing literature of synchromodality, along with the applications of external cost on freight transportation. In section 3, a mathematical model for the synchromodal transportation problem is presented. The section 4 introduces the NSGA-II algorithm, which is employed as a metaheuristic method to solve the problem. The model validated on generated instances and the computational results are discussed in section 5. The section 6 explores the impact of incorporating external costs, followed by the conclusion.

2 Literature Review

2.1 Modelling approaches in Synchromodal Transportation

Although synchromodal transportation problem is a relatively new paradigm, several modeling and solution methodologies have been developed to address its complexities. For instance, [11] formulated the synchromodal transportation problem as a k-shortest path problem, aiming to minimize total monetary costs, time delays, and carbon emissions in freight transport. The study demonstrated a reduction in average monetary costs by 10.1% and CO_2 emissions by 14.2%, primarily through a modal shift from road to rail. A link-based modeling approach is used in [1, 4], where the transportation network is represented as a graph, with intermodal terminals as nodes and transportation services modeled as arcs. Additionally, methods such as approximate dynamic programming (ADP) have been applied to address stochastic elements in synchromodal transportation, as demonstrated by [3].
Solving the synchromodal transportation problem on large networks—both in static and dynamic settings—often requires the use of heuristic or metaheuristic techniques to find solutions within a reasonable time frame. For example, a pre-processing path generation heuristic for shipment-matching in synchromodal transportation was proposed in [16], focusing on minimizing transportation costs and carbon emissions. Similarly, [17] implemented an Adaptive Large Neighborhood Search (ALNS) to address a synchromodal transportation problem that incorporated vehicle routing, allowing for flexible origins and destinations for vehicles. The objectives included minimizing transportation costs, transfer costs, storage charges, carbon taxes, waiting costs, and delay penalties. In another study [9], authors integrated vehicle routing problems into an intermodal service network design problem, and used a sample average approximation combined with iterated local search to solve the problem efficiently.

A multi-objective Genetic Algorithm (GA) using a weighted sum approach was used as the solution methodology for a synchromodal transportation problem in [15], where the objectives were to minimize total transportation costs, transshipment costs, and delay and overdue penalties. Similarly, [7] introduced a bi-level multi-objective Taguchi GA for multimodal routing, aimed at minimizing transportation time and travel costs. Their solution generated a Pareto front of optimal routes, enabling shippers to choose based on their preferences. Lastly, [18] proposed a preference-based multi-objective optimization model for synchromodal transportation, formulated as a pickup and delivery problem with transshipment. This model considered three objectives: minimizing costs, emissions, and time. An ALNS metaheuristic using a weight interval method was employed to produce a Pareto frontier of optimal solutions.

2.2 External Costs in Multimodal Transport

The European Union has been actively working to internalize external costs in transport, aiming to evaluate and monetize these costs for integration into cost estimation frameworks. A comprehensive *Handbook on External Costs* [6], published in 2019 provides state-of-the-art estimates for all major external transport costs. In this study, we refer to this handbook for the average external costs values associated with various modes of transportation.

According to the handbook, the external costs can be categorized into several categories:

- Accident Costs: These encompass both material (e.g., vehicle damage) and immaterial (e.g., reduced life expectancy) expenses resulting from transport accidents.
- Air Pollution Costs: Arising from emissions of transport-related air pollutants such as particulate matter (PM_{10}, $PM_{2.5}$), nitrogen oxides (NO_x), and sulfur dioxide (SO_2), these costs contribute to health issues, crop failures, infrastructure degradation, and environmental damage.
- Climate Change Costs: Related to air pollution but specifically driven by greenhouse gases like CO_2, N_2O, and CH_4, these costs contribute to global warming and rising temperatures.
- Noise Costs: Referring to unwanted noise of varying duration or intensity, these costs can cause physical or psychological harm.
- Congestion Costs: These represent the delays and welfare losses caused by traffic congestion, particularly in road transport.
- Well-to-Tank (WTT) Emissions: These include the environmental costs associated with energy production for transportation, covering processes such as resource extraction, processing, and transmission.
- Habitat Damage: This represents costs arising from the negative environmental impacts of transportation, including ecosystem loss and habitat degradation.

The average external costs (in Euro/tonne.km.) for each mode of transport, categorized by externality types, are illustrated in Figure (1). The figure provides a breakdown of external costs across various transport modes, including road, rail (diesel and electric), and inland waterway transport (IWT). Each mode exhibits different cost intensities across the above-mentioned categories. For instance, road transport incurs significantly higher external costs due to its prominent impact on congestion, accidents, and emissions. In contrast, rail and IWT are generally more sustainable, with lower external costs related to accidents, air pollution, and congestion, but their external costs in categories like WTT emissions and habitat damage are still notable. The process of incorporation of external costs in the cost analysis and decision-making, in transportation is known as internalization.

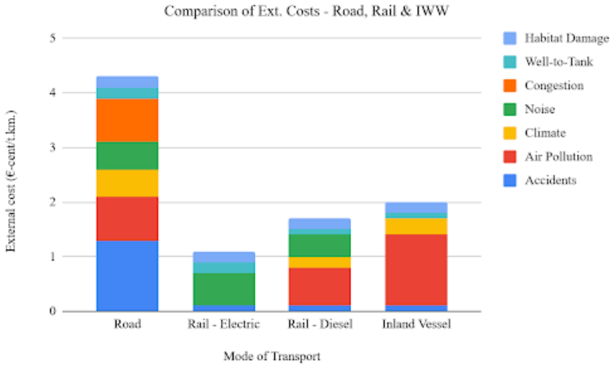


FIG. 1: External costs for various modes of transport

In the study [12], authors demonstrated that internalizing external costs could enhance the market share of intermodal transport, as rail and IWT tend to be more sustainable alternatives to road-based transport. Minimizing the external costs by employing more efficient vehicles rather than solely promoting modal shifts—can also help reduce emissions and mitigate the negative externalities associated with freight transport. For instance, in [8], authors proposed a numerical model to evaluate emissions in the intermodal transportation chain, accounting for both internal and external costs. Their findings indicated that utilizing IWT powered by sustainable fuels, such as low-sulfur alternative fuel, could decrease external costs, enhancing the competitiveness of IWT with road transport. However, they found that combining monetized external and internal costs does not significantly affect transportation chain preferences, as external costs represent only 26% of internal costs and thus have limited influence.

[14] highlighted that internalizing external costs may not always be advantageous, especially concerning the pre- and post-haulage components of intermodal transport. Nonetheless, future transportation systems powered by alternative fuels, such as hybrid or electric vehicles, are likely to positively impact both external and internal costs by reducing emissions and consumption. Increasing the subsidies that encompass operating, external, and transfer costs for rail and IWT, along with higher carbon taxes and external cost coefficients for road transport was advocated by [5]. This combination is expected to encourage a modal shift from road transportation to more sustainable alternatives such as IWT and Rail. A comprehensive review of external costs in road and intermodal transportation, including various external cost categories and their associated objectives is given in [13].

In this study, we propose an NSGA-II algorithm that eliminates the need for normalization of objectives. The transportation network is modeled as a multigraph, utilizing a novel encoding technique that directly represents the parallel edges of the graph. Additionally, this study explicitly internalizes external costs, an often-overlooked factor in synchromodal transportation problems.

3 Problem Statement

The synchromodal transportation problem aims to find feasible transportation plans between an origin and a destination location for container transportation while minimizing the total carbon emissions, total travel time and total external costs. The mathematical model presented in the work [2] is extended to consider external costs.

The synchromodal network is represented as a graph $G = (N, S)$, where N denotes the set of nodes, acting as intermodal terminals, and S represents the set of edges, acting as transportation services connecting terminal pairs $(i, j) \in N$. In this work, three transport modes are considered: IWT, rail, and road. Rail services are further subdivided into electric and diesel-powered trains, distinguished by their respective average external cost factors. Each transportation service $s \in S$ is characterized by attributes such as service ID, earliest and latest start times, duration, distance, capacity, carbon factor, and average external cost factor.

Customer orders, referred to as transportation requests, are denoted by the set R, with each request $r \in R$ specifying the origin, destination, number of containers, and weight. The model incorporates practical constraints, including limited capacity and fixed schedules for IWT and rail, while trucks are assumed to be available on demand without capacity limitations. Terminal operating hours and a fixed average transshipment time are also taken into account.

Parameters

The model has following parameters.

- D_{ij}^{s} : Duration of the service s between terminals i and j.
- $Dist_{ij}^{s}$: Distance between terminals i and j using service s.
- EC_s : External Cost factor for a service s.
- Q_s : Capacity of a service s.
- q^r : Units of containers or demand corresponding to a request r.
- $\bar{o}$, $\bar{d}$: Origin and Destination of the request.

- τ^r_{rel} : Release time of a request
- τ^r_{due} : Due time of a request.
- TO_i : Opening hour of terminal at i.
- TC_i : Closing hour of terminal at i.
- TS_i : Transshipment time per unit container at terminal at i.
- α^s : Earliest starting time of a service.
- β^s : Latest starting time of a service.
- cw_s : Carbon emission factor of a service.
- θ_r : Weight of the TEUs of a request.
- M : Large enough number

Decision Variables

- K^s_{ij} : Binary variable, equals to 1 if a service s is used between the terminals i and j, 0 otherwise.
- X^{sr}_{ij} : Non-negative variable representing the number of containers for request r served by service s between terminals i and j.
- Y^{sr}_{ij} : Binary variable, equals to 1 if a service s is used between the terminals i and j to serve the request r, 0 otherwise.
- Z^{sl} : Binary variable, equals to 1 if there is a precedence between two services s and l, 0 otherwise.
- ST^r_i : Storage time of request r at terminal i.
- T^s_{ij} : Starting time of service s between i and j.
- A^s_{ij} : Arrival time of service s between i and j.

Objective Function
The objective functions considered to be minimized in this study are:

- Carbon Emissions:
$$f_1 = \sum_{i,j \in N} \sum_{s \in S} \sum_{r \in R} \theta_r \cdot cw_s \cdot Dist^s_{ij} \cdot Y^{sr}_{ij}$$
- Transportation Time:
$$f_2 = \sum_{i,j \in N} \sum_{s \in S} \sum_{r \in R} D^s_{ij} \cdot Y^{sr}_{ij}$$
- External Costs:
$$f_3 = \sum_{i,j \in N} \sum_{s \in S} \sum_{r \in R} EC_s \cdot \theta_r \cdot Dist^s_{ij} \cdot Y^{sr}_{ij}$$

$$\min\left(f_1, f_2, f_3\right) \tag{1}$$

Constraints
subject to.

$$\sum_{j \in J} K^s_{oj} \leq 1, \quad \forall s \in S \tag{2}$$

$$\sum_{j \in N} K^{s}_{\bar{o}j} = \sum_{j \in N} K^{s}_{j\bar{d}}, \quad \forall s \in S \tag{3}$$

$$\sum_{j \in N} K^{s}_{ij} - \sum_{j \in N} K^{s}_{ji} = 0, \quad \forall s \in S, \quad \forall i \in N \setminus \bar{o}, \bar{d} \tag{4}$$

$$\sum_{s \in S} X^{sr}_{ij} - \sum_{s \in S} X^{sr}_{ji} = \begin{cases} q^{r}, \quad \textit{if } i = \bar{o}, \quad \forall r \in R, \quad \forall i, j \in N \\ -q^{r}, \quad \textit{if } i = \bar{d}, \quad \forall r \in R, \quad \forall i, j \in N \\ 0, \quad \textit{otherwise} \end{cases} \tag{5}$$

$$\sum_{r \in R} X^{sr}_{ij} \leq M \cdot Y^{sr}_{ij}, \quad \forall s \in S, \quad \forall r \in R, \quad \forall i, j \in N \tag{6}$$

$$K^{s}_{ij} \geq Y^{sr}_{ij}, \quad \forall s \in S, \quad \forall r \in R, \quad \forall i, j \in N \tag{7}$$

$$\sum_{r \in R} X^{sr}_{ij} \leq Q_{s} \cdot K^{s}_{ij}, \quad \forall s \in S, \quad \forall i, j \in N \tag{8}$$

$$T^{s}_{ij} + D^{s}_{ij} - A^{s}_{ij} \leq M \cdot (1 - K^{s}_{ij}), \quad \forall s \in S, \quad \forall i, j \in N \tag{9}$$

$$\begin{aligned} & A^{s}_{ij} + q^{r} \cdot TS_{j} + ST^{r}_{j} - T^{l}_{jk} \leq M \cdot (3 - Y^{sr}_{ij} - Y^{sr}_{jk} - Z^{sl}), \\ & \forall i, j, k \in N, \quad \forall s, l \in S, \quad \forall r \in R \end{aligned} \tag{10}$$

$$\begin{aligned} & A^{s}_{ij} + q^{r} \cdot TS_{j} - ST^{l}_{jk} \leq M \cdot (1 - Z^{sl}), \\ & \forall s, l \in S, \quad \forall r \in R, \quad \forall i, j, k \in N \end{aligned} \tag{11}$$

$$Z^{sl} + Z^{lq} + Z^{qs} <= 2, \quad \forall s, l, q \in S \tag{12}$$

$$Z^{sl} + Z^{ls} <= 1, \quad \forall s, l \in S \tag{13}$$

$$\alpha^{s} \leq T^{s}_{ij} \leq \beta^{s} \quad \forall s \in S, \quad \forall i, j \in N \tag{14}$$

$$TO_{i} \leq T^{s}_{ij} \leq TC_{i} \quad \forall s \in S, \quad \forall i, j \in N \tag{15}$$

$$\tau^{r}_{rel} \leq T^{s}_{\bar{o}j} \quad \forall s \in S, \quad \forall r \in R \quad \forall j \in N \tag{16}$$

$$A^{s}_{i\bar{d}} + q^{r} \cdot TS_{\bar{d}} \leq \tau^{r}_{due} \quad \forall s \in S, \quad \forall r \in R \quad \forall i \in N \tag{17}$$

$$Y^{sr}_{ij} + Y^{sr}_{jk} + Y^{sr}_{ki} \leq 2, \quad \forall s \in S, \quad \forall r \in R, \quad \forall i, j, k \in N \tag{18}$$

$$ST^{r}_{i}, T^{s}_{ij}, X^{sr}_{ij} \geq 0, \quad \forall s \in S, \quad \forall r \in R, \quad \forall i, j \in N \tag{19}$$

$$K^{s}_{ij}, Y^{sr}_{ij}, Z^{sl} = \{0, 1\}, \quad \forall s, l \in S, \quad \forall r \in R, \quad \forall i, j \in N \tag{20}$$

The objective function (1) minimizes total carbon emissions, transportation time and external costs. Constraints (2) and (3) ensure services start from their origins and end at their destinations. Constraint (4) enforces flow balance. Constraint (5) ensures transportation of each request. Constraints (6)-(8) link requests to selected services. Constraint (9) respects service capacities. Constraint (10) links service arrival times to starting times and travel durations. Constraint (11) maintains time continuity with transshipment and storage. Constraints (12-13) enforce precedence when one service arrives before another departs. Constraints (14) and (15) ensure time window and terminal working hours adherence. Constraints (16) and (17) respect order release and delivery times. Constraint (18) prevents sub-tours. Constraints (19)-(20) are domain constraints.

Since it is practically infeasible to solve large-sized instances in respectable timeframe using exact methods, a NSGA-II algorithm is proposed in the following section to solve large instances of synchromodal transportation problem.

4 NSGA-II algorithm for Synchromodal Transportation Problem

NSGA-II, a population-based metaheuristic introduced by [10], is widely used for solving combinatorial optimization problems. As an improvement over its predecessor, NSGA, it incorporates two key mechanisms: fast non-dominated sorting and crowding distance assignment. The non-dominated sorting ranks solutions based on their dominance relationships with other solutions in the population, giving priority to non-dominated solutions. The crowding distance, on the other hand, measures how densely solutions are clustered, ensuring diversity is maintained within the Pareto front. The general structure of the NSGA II algorithm is given in Algorithm 1.

Algorithm 1 NSGA-II Algorithm

```
Require: $\eta$, $maxG$, $Prob_c$, $Prob_m$, $R$, $S$
Ensure: $NSGA-II$
while $R$ is not empty do
    Pull the first request from $R$
    Perform the following steps for the current request:
GA Steps:
    Initialize generation counter $gen = 1$
    Generate $\eta$ solutions to form the initial population, $P$
    Evaluate the fitness of solutions in $P$
    Assign ranks based on non-dominated sorting
    Assign crowding distance for each solution
    while $gen \leq maxG$ do
        Perform crossover with probability $Prob_c$:
                Choose two solutions $p_1$ and $p_2$ from $P$.
                Generate two children $c_1$ and $c_2$ using crossover operator
        Mutate $c_1$ and $c_2$ with probability $Prob_m$
        if $c_1$ and $c_2$ are not in population $P$ then
            Insert offspring $c_1$ and $c_2$ into population $P$
        end if
        Perform non-dominated sorting on combined population $P$
        Select the best $N$ solutions to form the new population
        if stopping criterion is satisfied or $gen < maxG$ then
            Return non-dominated solutions as $bestSolutions$
        end if
        Increment generation counter: $gen = gen + 1$
    end while
end while
```

The algorithm starts by taking the following parameters as input: Population size (η), the maximum number of generations ($maxG$), probability of crossover ($Prob_c$), probability of mutation ($Prob_m$), set of requests(R), set of services(S) and outputs a set of feasible transportation paths $bestSolutions$ for each request, $r \in R$.

The algorithm creates an initial population of size (η) using the path-generating operators given in section 4.2. A non-dominated sorting is then performed on the initial population to assign ranks to each solution in the population. After ranking the solution, the crowding distance procedure is performed for each solution. The aim is to prevent the convergence of the solution and explore extreme points of the solution space. In the next step, two-parent solutions (P_1) and (P_2) are selected using the tournament selection procedure. The crossover and mutation procedures are then performed on the parent solutions to generate offspring (C_1 and C_2). The new off-spring solutions are re-inserted into the population (P), assuming the off-spring are not existing in the population. This combined population undergoes another non-dominated ranking procedure and only the best η solutions are selected for the next iteration. The procedure is repeated until a maximum number of iterations ($maxG$) or a stopping criterion is reached. The stopping criterion is 50 generations without improvement in the non-dominating solutions. The algorithm ends by returning the non-dominated set of solutions $bestSolutions$ for each request $r \in R$.

4.1 Chromosome encoding

Chromosome encoding, which represents an individual solution, is one of the key aspects of NSGA-II algorithm. In this study, each solution is encoded as a transportation route from an origin terminal to a destination terminal. Given a synchromodal transportation network $G = (N, A)$, the chromosome begins with the origin terminal of the request (O_r) and ends in the destination terminal of the request (D_r). The genes at the odd position represent intermodal terminals (i.e. nodes N) and the genes in the even location represent the ID of the transportation service connecting the two intermodal terminals. The length of the chromosome is adaptive, based on the number of intermediate terminals and the maximum size will be $2N - 1$, assuming that there are N nodes with at least one transportation service between each node. A sample chromosome encoding is given in Figure (2), which represents a solution from an origin (o) to destination (d) using the services $[S_1, S_2, S_3]$ and intermediate terminals $[T_1, T_2]$:

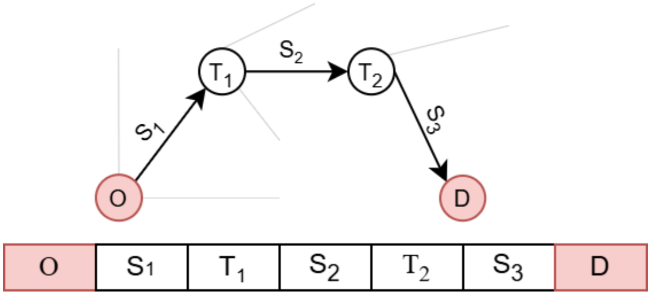


FIG. 2: Sample solution and chromosome encoding

4.2 Initial Population

The initial population is generated through three path generation operators: Direct path, Random path and One-stop path operators. Direct path operator finds a direct transportation service to transport the containers from request origin to destination and the random path finds a random travel route for transportation. The goal of one-stop path operator is to find a transportation service from request origin to an intermediate terminal and then another towards request destination. These three operators are iteratively applied until a fixed number (η) of non-duplicated paths is obtained.

4.3 Genetic Operators : Crossover & Mutation

The crossover operator is used to explore new areas in the solution space while the mutation operator helps in escaping a local optima by making small changes to the chromosome. In the NSGA II, the crossover is performed by selecting two parents with a common intermediate terminal, which will act as the crossover point. The parent selection is performed using tournament selection procedure and it is ensured that the selected parents will have at least one crossover point. Once the crossover point is found, the genes preceding the crossover point in parent 1 are combined with the genes succeeding the point in parent 2 and vice versa, to create two new children The crossover procedure is demonstrated in Figure (3), where the crossover point is shown in green color:

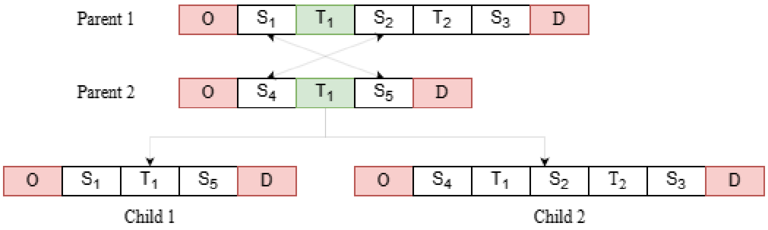


FIG. 3: Crossover Operator.

Mutation operator works by changing the individual genes in the chromosome. In the proposed algorithm, the mutation operator swaps a randomly chosen service between two terminals in the chromosome with an alternate service between the same terminal. A sample mutation operation is given in Figure (4), where the service in green color will be mutated. The off-springs generated by the crossover and mutation operators are ensured to be feasible with respect to capacity, time windows, and synchronization between preceding and succeeding services.

FIG. 4: Mutation Operator.

4.4 Parameter Tuning

In this study, a Taguchi analysis based on Pareto performance indicators is used to adjust the following NSGA-II parameters, with each having four levels: (a) Population Size [30, 40, 50, 60], (b) Number of Generations [100, 150, 200, 250], (c) Crossover Rate [0.3, 0.5, 0.7, 0.9], and (d) Mutation Rate [0.1, 0.2, 0.3, 0.4]. A L16 (4^4) Type B orthogonal design array is used for the analysis and the responses are the cumulative sum of 4 pareto performacnce indicators: (1) Number of Pareto Solutions (NPS), (2) Hypervolume, (3) Spacing, and (4) Best Solution.

NPS refers to the number of non-dominated solutions in the Pareto front, with higher values desirable for exploring a wide range of potential solutions. Hypervolume metric measures the volume of the objective space dominated by the non-dominated solutions in a Pareto front, with a larger hypervolume indicating broader solution space coverage. Spacing measures the Euclidean distance between solutions in the Pareto front, with smaller values indicating a more uniform distribution of solutions. Finally, the Best Solution represents the optimal values found by the algorithm, identified by selecting solutions with the lowest values for each objective and normalizing them to ensure comparability across different objectives. From the analysis, the identified optimal parameter combination is [60, 250, 0.7, 0.3], respectively. The results of the Taguchi analysis in graphical form is given in Figure (5).

5 Experimental Results

A set of instances are generated to validate the model and the algorithm. The instances are based on the Seine Axis in France and are differentiated into 4 groups: G1, G2, G3, and G4. Each group differs in size, with respect to the number of transportation services as well as the number of requests, with G1 being the smallest and G4 being the largest. A transportation service is characterized by an origin terminal, a destination terminal, mode of transport (IWT, rail or road), time schedules, capacity, and carbon emission factors. The capacity of barges is set at 300 TEUs, train at 90 TEUs, while trucks are assumed to have infinite capacity, reflecting the model's assumption that road transportation is always available. Additionally, each transportation service includes an average external cost, measured in Euro/tonne.km.,

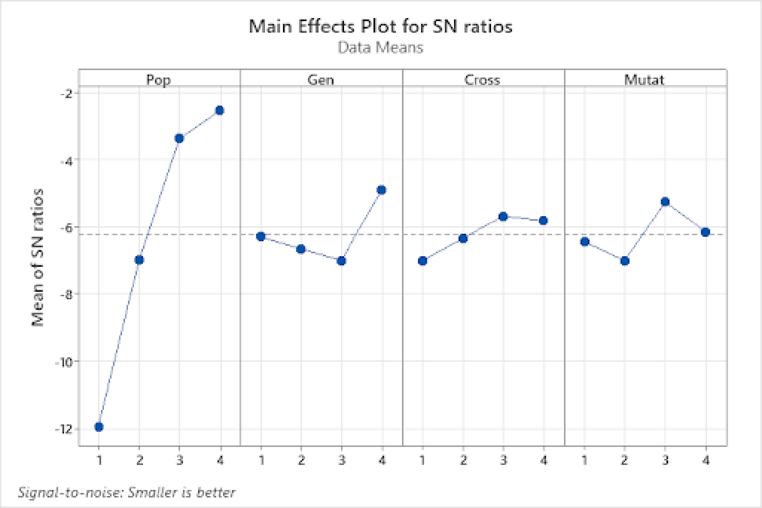


FIG. 5: Taguchi Analysis : S/N Ratio

which depends on the transportation mode.The mode share distribution in the instances is: 50% Road, 30% IWT, and 20% Rail, with trains being both diesel (10%) and electric (10%) powered. Similarly, a transportation request includes an origin, a destination, available time, due deadline, capacity demand, and the total weight of the shipment.

The instances are implemented on a system equipped with an Intel(r) Core i7-1185G7 processor and 32 gigabytes of RAM, running Windows 11 OS. Smaller instances are solved using the IBM CPLEX 12.1 mathematical solver, and the solutions obtained are compared with those generated by NSGA-II. For NSGA-II, a Pareto frontier is generated for each request within an instance. The Pareto front obtained for a sample request is illustrated in Figure (6).

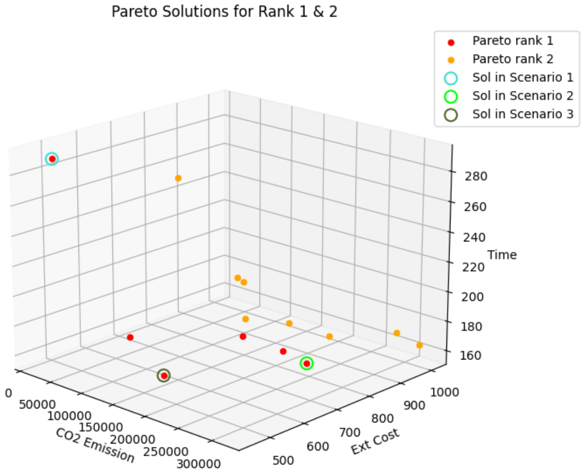


FIG. 6: Best Pareto solutions for scenarios.

The figure illustrates solutions belonging to the first and second pareto frontiers. As observed, multiple solutions exist in the Pareto frontier, providing flexibility to the user to choose a route based on their specific requirements. This approach aligns with the concept of mode-free booking in synchromodal transportation, offering a variety of transportation itineraries. Since each request is treated individually, a solution from the Pareto front must be selected as the preferred solution for the current request before moving on to the next request. In the

Scenario 1					
Instance	Obj_CO2 - MILP	CPU - MILP	Obj_CO2_NSGAII	CPU - NSGAII	Dev %
G1_1	52462.1	2.126	52462.1	0.1536	0.0%
G1_2	18580.1	1.894	18580.1	0.377	0.0%
G1_3	22231.6	2.759	22231.6	1.2903	0.0%
G1_4	93019.5	1.711	93019.5	0.1775	0.0%
G2_1	58670.4	72.9	59570.4	4.70884	1.53%
G2_2	54791.8	69.073	54791.8	4.4964	0.00%
G2_3	40100.2	67.48	40274.19	3.7705	0.43%
G2_4	122159.9	98.673	125041.4	4.823025	2.36%
Scenario 2					
Instance	Obj_Dur - MILP	CPU - MILP	Obj_Dur_NSGAII	CPU - NSGAII	Dev %
G1_1	362	2.031	362	0.220	0.00%
G1_2	396	2.022	396	0.144	0.00%
G1_3	323	2.835	323	0.354	0.00%
G1_4	382	2.071	382	0.115	0.00%
G2_1	447	38.362	447	4.094	0.00%
G2_2	654	52.036	654	4.778	0.00%
G2_3	428	46.203	435	4.502	1.64%
G2_4	570	52.173	570	5.087	0.00%
Scenario 3					
Instance	Obj_EC - MILP	CPU - MILP	Obj_EC _NSGAII	CPU - NSGAII	Dev %
G1_1	1693.79	5.889	1693.79	0.151	0.00%
G1_2	1790.90	4.946	1790.90	0.358	0.00%
G1_3	1377.10	4.138	1377.10	0.264	0.00%
G1_4	2027.52	3.660	2027.52	0.126	0.00%
G2_1	2313.24	95.746	2318.79	5.618	2.40%
G2_2	3513.80	94.178	3513.80	5.422	0.00%
G2_3	1939.44	115.427	1943.12	5.915	1.90%
G2_4	2930.00	91.163	2938.79	5.567	3.00%

TAB. 1: Results for instance group 1 and 2 - MILP vs NSGA-II

algorithm, three scenarios are designed for this purpose:

- Scenario 1: The solution with the lowest carbon emissions is selected as the preferred solution.
- Scenario 2: The solution with the shortest travel duration is chosen.
- Scenario 3: The solution with the lowest external costs is selected.

For the illustrated request, the solution with the lowest value in each scenario is shown in figure 6, where the preferred solutions are marked with a colored circle.

The solutions obtained from NSGA-II is compared against the solutions obtained from the mathematical solver. Each scenarios were individually solved, and it was observed that NSGA-II provides solutions that are on par with the mathematical model. In the smallest instance group (G1), NSGA-II matches the optimal solutions found by the MILP solver in almost all cases, with a low deviation of less than 5% in some instances from group G2. The comparison of results obtained from solver and NSGA-II algorithm is given in Table 1.

As instance sizes increase, the computational time for the mathematical solver rises exponentially, and it struggles with larger instances, often hitting computational limits. In contrast, NSGA-II provides optimal or near-optimal solutions in all cases, while being significantly faster than the MILP solver. The detailed results, including large instance sizes, are provided in the Table 3. The objectives represent the average solutions from 10 runs of the NSGA-II algorithm. The deviation from the best-found solution is calculated using the following formula:

Mode	Road	Rail	IWT
Scenario 1	20.00%	6.60%	73.40%
Scenario 2	69.23%	30.77%	0.00%
Scenario 3	50.00%	33.40%	16.60%

TAB. 2: Mode Share - Road vs Rail vs IWT

$$\text{Deviation (\%)} = \frac{\text{Average Solution} - \text{Best Solution}}{\text{Best Solution}} \times 100$$

In all cases, the deviation remains below 5%, indicating consistent solution quality across runs.

6 Impact of External Costs on Mode Share

The mode share percentage of each transportation mode, in each scenario is given in Table 2. From the table, it can be observed that the utilization of rail transportation is increased while minimizing external costs in scenario 3. This can result in a reduced travel duration, when compared to scenario 1 and a reduced total carbon emission, when compared to scenario 2. Additionally, the usage of IWT is significantly lower when the external costs are minimized. This can be attributed to the long distances that IWT has to cover, resulting in a higher external cost. Hence, it can be concluded that internalization of external costs can increase the utilization of rail and road transportation, while reducing the IWT share.

	Scenario 1					Scenario 2					Scenario 3				
Instance	Obj_CO2	Obj_Duration	Obj_Ext_Cost	CPU	Dev	Obj_CO2	Obj_Duration	Obj_Ext_Cost	CPU	Dev	Obj_CO2	Obj_Duration	Obj_Ext_Cost	CPU	Dev
G1_1	52462.1	725	1693.78	0.154	0.00%	106971.2	362	1900.11	0.220	0.00%	52462.10	725	1693.79	0.151	0.00%
G1_2	18580.1	992	1812.99	0.377	0.00%	129930	396	2302.50	0.144	0.00%	18663.10	1005	1790.90	0.358	0.00%
G1_3	22231.6	1091	1868.5	1.290	0.00%	79687.9	323	1576.84	0.354	0.00%	30296.30	640	1377.10	0.264	0.00%
G1_4	93019.5	636	2248.45	0.178	0.00%	113485.8	382	2132.66	0.115	0.00%	103849.50	383	2027.52	0.126	0.00%
G2_1	59570.4	1422	3031.4291	4.709	0.00%	136160.5	447	2523.66	4.094	0.00%	84741.90	758	2313.24	5.618	0.00%
G2_2	54791.8	1999	3719.28	4.496	1.30%	241095.2	654	4098.40	4.778	0.00%	69613.40	1756	3513.80	5.422	0.05%
G2_3	40274.19	1459	2793.95	3.771	1.10%	109844.8	435	2213.03	4.502	0.00%	62837.20	586	1939.44	5.915	2.12%
G2_4	125041.4	934	3164	4.823	0.00%	156880.6	570	2995.70	5.087	0.00%	140862.00	666	2930.00	5.567	0.03%
G3_1	1300681.09	10579	19243.8	17.227	1.67%	7964460.17	5412	26849.27	15.902	0.47%	1889344.51	8743	17620.27	17.970	0.58%
G3_2	1781165.33	9642	18984	17.760	1.61%	7487609.44	5801	26205.69	15.958	0.71%	2332060.33	7810	17302.25	18.661	0.60%
G3_3	1632661.17	9789	18777.02	13.452	2.52%	7160363.36	5118	24549.76	15.315	1.00%	1806816.36	9003	17884.91	15.236	3.05%
G3_4	1016954.32	10036	17625.84	12.023	1.25%	6967867.18	4944	24285.03	17.008	0.48%	1349245.95	8644	16176.6	16.089	0.44%
G4_1	1622239.73	22859	37971.63	51.420	1.61%	14766508.79	10406	50579.34	51.700	2.66%	2489284.33	20444	36242.13	51.003	1.17%
G4_2	2043823.52	23322	40106.74	51.610	0.05%	14273244.79	10563	49289.24	51.357	2.93%	3363793.17	19198	36004.07	52.302	1.99%
G4_3	2666153.82	23460	42057.36	49.930	0.98%	15446014.54	10231	52081.18	50.880	3.92%	4367255.80	18646	37526.39	51.475	0.83%
G4_3	2207309.91	23415	40304.81	52.090	0.45%	15639243.9	11036	53814.05	51.730	1.64%	3578388.16	19918	37850.27	50.660	0.54%

TAB. 3: Detailed results - NSGA-II

7 Conclusion

The study focuses on the multi objective optimization of the synchromodal transportation problem. A mathematical model is introduces which considers three modes of transport: IWT, rail, and road. To solve instances of larger sizes, a NSGA-II multi-objective metaheuristic is proposed. A novel encoding method along with crossover and mutation operators for NSGA-II are presented. The algorithm's parameters are determined through Taguchi's method, by using Pareto performance indicators as the response variable. A set of instances based on the Seine Axis in France are generated to validate the model.

Smaller sized instances are solved using a MILP solver while larger instances reaches computational limit. NSGA-II on the other hand is able to solve instances of all sizes rapidly with a good solution quality. The study considers internalization of external costs along with total carbon emissions and duration in the objectives, and the effect of each objective on the mode choices are observed. When external costs are considered in mode choice, it can be observed that the utilization of rail mode is higher while the utilization of IWT is reduced. In the future, this study can be extended by integrating additional objectives, such as user preferences, into the optimization framework, thereby enhancing the decision-making process in synchromodal transportation.

References

[1] Aaron Agbo and Yanwei Zhang Sustainable freight transport optimisation through synchromodal networks. *Cogent Engineering*, 4(1) 2018.

[2] Ananthakrishnan Vaikkathe, Abdelhamid Benaini, and Jaouad Boukachour. A genetic algorithm for synchromodal transportation problem. *Proceeding of the 7th International Conference on Logistics Operations Management, GOL'24*, pages 191–201. Springer Nature Switzerland, Cham, 2024.

[3] Arturo Pérez Rivera and Martijn Mes. Service and transfer selection for freights in a synchromodal network. *Computational Logistics*, 227–242. 2016.

[4] Behzad Behdani, Yun Fan, Bart Wiegmans, and Rob Zuidwijk. Multimodal schedule design for synchromodal freight transport systems. *EJTIR Issue* 16, 424–444 2016.

[5] David Austin. Pricing overland freight transport to account for external costs. *Journal of Transport Economics and Policy*, 52(1):45–67, 2018..

[6] European Commission: Directorate-General for Mobility and Transport, H. Essen, D. Fiorello, K. El Beyrouty, C. Bieler, L. Wijngaarden, A. Schroten, R. Parolin, M. Brambilla, D. Sutter, S. Maffii, and F. Fermi. *Handbook on the external costs of transport – Version 2019 – 1.1.* Publications Office, 2020.

[7] Guiwu Xiong and Yong Wang. Best routes selection in multimodal networks using multi-objective genetic algorithm. *Journal of Combinatorial Optimization*, 28(3):655–673, 2014.

[8] Heitor Abreu, Tiago A. Santos, and Valdir Cardoso. Impact of external cost internalization on short sea shipping – The case of the Portugal-Northern Europe trade. *Transportation Research Part D: Transport and Environment*, 114:103544, 2023.

[9] Jan Philipp Müller, Ralf Elbert, and Simon Emde. Integrating vehicle routing into intermodal service network design with stochastic transit times. *EURO Journal on Transportation and Logistics*, 10:100046, 2021.

[10] Kalyanmoy Deb, Amrit ratap, Sameer Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[11] Martijn R.K. Mes and Mihnea E. Iacob. Synchromodal transport planning at a logistics service provider. *Logistics and Supply Chain Innovation: Bridging the Gap between Theory and Practice*, 23–36. Springer, 2016. (Lecture Notes in Logistics).

[12] Martin Mostert, An Caris, and Sabine Louise Limbourg. Road and intermodal transport performance: The impact of operational costs and air pollution external costs. *Research in Transportation Business & Management*, 23:75–85, 2017.

[13] Martine Mostert and Sabine Limbourg. External costs as competitiveness factors for freight transport – A state of the art. *Transport Reviews*, 36(6):692–712, 2016.

[14] Mónica M. Ramalho and Tiago A. Santos. The impact of the internalization of external costs in the competitiveness of short sea shipping. *Journal of Marine Science and Engineering*, 9(9):959, 2021.

[15] Nasibeh Zanjirani Farahani, James S. Noble, Ronald G. McGarvey, and Moein Enayati. An advanced intermodal service network model for a practical transition to synchromodal transport in the US Freight System: A case study. *Multimodal Transportation*, 2(1):100051, 2023.

[16] Wenjing Guo, Bilge Atasoy, Wouter Beelaerts van Blokland, and Rudy R. Negenborn. A dynamic shipment matching problem in hinterland synchromodal transportation. *Decision Support Systems*, 134:113289, 2020.

[17] Yimeng Zhang, Wenjing Guo, Rudy R. Negenborn, and Bilge Atasoy. Synchromodal transport planning with flexible services: Mathematical model and heuristic algorithm. *Transportation Research Part C: Emerging Technologies*, 140:103711, 2022.

[18] Y. Zhang, B. Atasoy, and R. R. Negenborn. Preference-based multi-objective optimization for synchromodal transport using adaptive large neighborhood search. *Transportation Research Record*, 2676(3):71–87, 2022.

Applications of Educational Data Mining in Handling Big Data: A Systematic Literature Review

AFIF Fatima-ezzahra [1], BOUYAHIA Fatima [2], RAFOUK Leila [1]

[1] Cadi Ayyad University, Laboratory of Didactic and University Pedagogy, Morocco
fatimaezzahra.afif02@gmail.com, rafouk@uca.ac.ma
[2] Cadi Ayyad University, Laboratory of Engineering and Systems Analysis, Morocco
f.bouyahia@uca.ac.ma

Abstract. The world is changing at rapid rhythm, the rise of digital technologies such as big data, artificial intelligence, simulation... is creating drastic innovations and new opportunities for learning and development. Big data certainly has become a game-changer in many industries, and education is not an exception. The large data sets collection and analysis have provided the basis for new approaches in student monitoring, students' performance prediction, and improvement of teaching and learning processes. This data is mined and analyzed through Educational Data Mining (EDM), a cross-disciplinary field involving computer science, education, and statistics. It analyses and mines educational data to better understand students and their learning settings.

Keywords: *educational big data; educational data mining; student's performance*

1 Introduction

The major developments made by technology in the present world in connection to different fields show that change is taking place in every aspect of life, especially in the field of education. The increasing use of digital technologies in teaching and learning has led to the production of massive datasets in the educational environment, commonly known as educational big data [1]. This huge and diverse data landscape presents both opportunities and challenges for educators and researchers focused on enhancing educational outcomes [2].

To make the best use of this educational big data, EDM has thus become a critical approach for analyzing and interpreting such voluminous data [3]. This paper focuses on educational big data and EDM.

To perform the systematic review, the Kitchenham methodology is used [4] as shown in Figure 1. The following research questions (RQ) are crucial to its successful implementation:

RQ1: What are the key research areas in educational big data?

RQ2: Which methods and application areas are most commonly used in EDM?

RQ3: What are the limitations of educational big data and EDM, and what can be the further development?

The remainder of this paper is organized as follows: First, we describe the review methodology. Second, we present the definition and background of educational big data and its relevance. Then, we provide an overview of the major components of EDM and its growth. Moreover, we present the use of data mining methods in the resolution or analysis of problems that occur in educational settings. Finally, we address the limitations and challenges of adopting EDM, and provide insights for future research directions to enhance the interpretability and transparency of educational technologies.

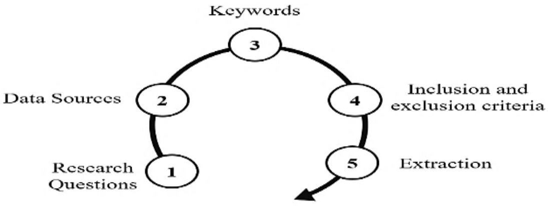


FIG. 1. Systematized overview methodology steps

2 Methodology

This section presents the Kitchenham methodology. As shown in Figure 2 below, the process began with a search ad extraction of titles, resulting in a total of 102 initial articles: 25 articles refer to educational big data and 77 are associated to EDM. During the second step, after evaluating the titles of the studies, 10 educational big data studies were excluded and, therefore, 15 studies were considered to proceed for further examination. For EDM, 20 articles were excluded and the remaining count was 57 for the next phase. The last stage involved classification based on the inclusion and exclusion criterion of abstracts, introductions, and the results. In the case of educational big data, a total of 9 papers are removed and 6 papers are identified for detailed examination. For EDM, 23 articles were removed, and 34 articles were included in this review. The majority of the reviewed works were published in 2020 (n=8), 2021 (n=8) and 2022 (n=7). The years 2023, 2024, and 2018 each have 4 studies included, while 2019 has 2 studies. The year 2017 includes 3 studies.

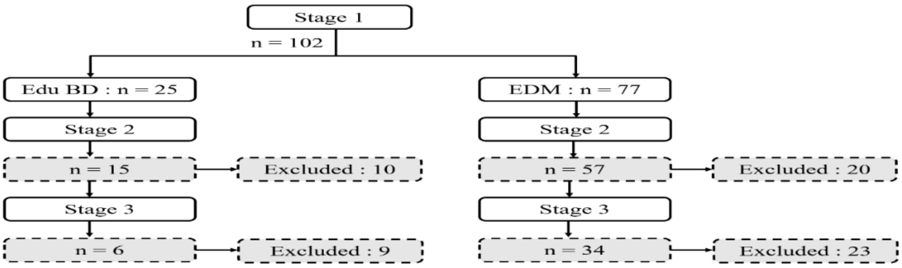
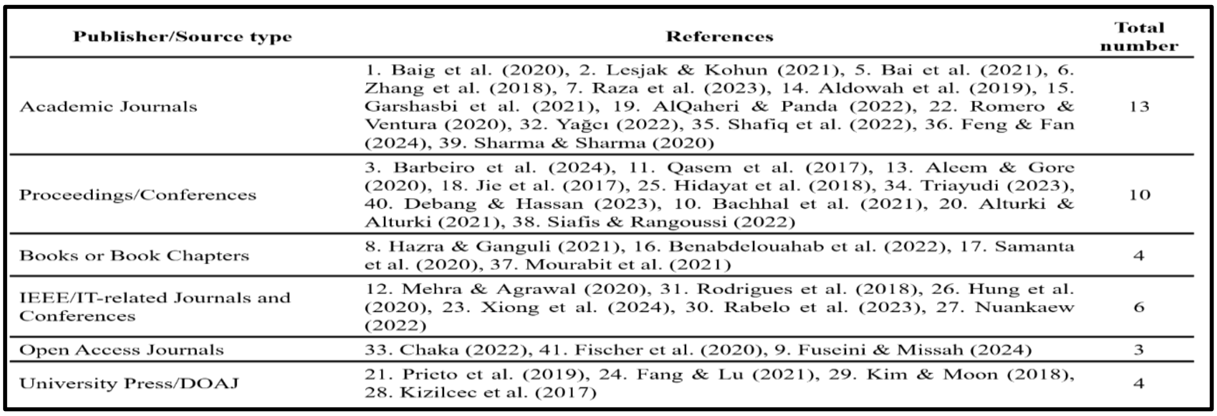


FIG 2: Process of selection studies

The table below provides an overview of the references of the publications on which this review is based according to their main source. The types of sources include academic journals, conference papers, books or book chapters, IEEE/IT-related journals and conferences, open-access journals, and university publishers /DOAJ (Directory of Open Access Journals). Here is a summary of the distribution:



Publisher/Source type	References	Total number
Academic Journals	1. Baig et al. (2020), 2. Lesjak & Kohun (2021), 5. Bai et al. (2021), 6. Zhang et al. (2018), 7. Raza et al. (2023), 14. Aldowah et al. (2019), 15. Garshasbi et al. (2021), 19. AlQaheri & Panda (2022), 22. Romero & Ventura (2020), 32. Yağcı (2022), 35. Shafiq et al. (2022), 36. Feng & Fan (2024), 39. Sharma & Sharma (2020)	13
Proceedings/Conferences	3. Barbeiro et al. (2024), 11. Qasem et al. (2017), 13. Aleem & Gore (2020), 18. Jie et al. (2017), 25. Hidayat et al. (2018), 34. Triayudi (2023), 40. Debang & Hassan (2023), 10. Bachhal et al. (2021), 20. Alturki & Alturki (2021), 38. Siafis & Rangoussi (2022)	10
Books or Book Chapters	8. Hazra & Ganguli (2021), 16. Benabdelouahab et al. (2022), 17. Samanta et al. (2020), 37. Mourabit et al. (2021)	4
IEEE/IT-related Journals and Conferences	12. Mehra & Agrawal (2020), 31. Rodrigues et al. (2018), 26. Hung et al. (2020), 23. Xiong et al. (2024), 30. Rabelo et al. (2023), 27. Nuankaew (2022)	6
Open Access Journals	33. Chaka (2022), 41. Fischer et al. (2020), 9. Fuseini & Missah (2024)	3
University Press/DOAJ	21. Prieto et al. (2019), 24. Fang & Lu (2021), 29. Kim & Moon (2018), 28. Kizilcec et al. (2017)	4

TAB. 1: References used according to their source type

3 Educational Big Data

3.1 5V's of Educational Big Data

Over the years, big data has revolutionized the educational systems in numerous ways. This term refers to the large and varied amounts of data that educational institutions gather, process, and apply to enhance not only learning and instructional strategies, but also operations, performance, and student outcomes. The emergence of educational data has brought both opportunities and challenges to the field of educational big data, which is characterized by five main characteristics: Volume, Velocity, Veracity, Variety, and Value [5]. Schools as learning institutions all over the globe produce a substantial amount of educational data which is part of the number aspect. The growth rate refers to the velocity aspect. Education big data contains many entities (students, teachers, administrators) and connections (teacher-student, classmate, friend) that make educational big data systems challenging due to such factors as name ambiguity (teachers, students) and data dependence. To overcome these challenges, however, it is vital to ensure the veracity of educational big data. The different characteristics of educational big data include behavioral aspects (fatigue, concentration, joy, surprise), life events (shopping, activities), personal information (genre, age, ethnicity, birth date, language, province, family), and learning behaviors (fraction completed, fraction spent, fraction played and fraction pause) [6]. These are just a few examples. The variety of features of educational platforms such as intelligent tutorial systems and MOOC platforms, tutorial transcripts, and reading materials include students' click-stream data [7].

The value of educational big data is one of its key features which is understanding the behaviour or pattern generated by the data. This leads to the next step of generating predictive models that will assist in identifying struggling learners and boosting teaching approaches to achieve improved education results. Educational institutions can use these patterns to make conclusions that would enhance the positive outcomes of the learning process as well as the efficiency of the process.

3.2 Educational Big Data Applications

Big data in the context of education provides insights to enhance learning, teaching, and assessment approach as compared to traditional methods. This section explores the most beneficial big data applications [5] [7] [8]:

A. Enhanced learning experience: Learning big data focuses on the educator and learner's needs with an attempt to optimize the teaching and learning environment to meet the learner and educator's resources and capability. Through accumulating large-scale student data and applying data mining techniques, models can be constructed to develop a personalized education delivery environment and provide recommendations for further support, which enhances opportunity and success rates in education.

B. Improved quality of education: In educational big data, decisions are made based on research and analysis, where decision-makers rely on insights derived from the appropriate datasets. This is especially concerning the efficient achievement of learning outcomes in schools to evaluate the learners' learning success rates, trends, challenges and performance. During the learning process, record is kept and documented the inputs and behaviors of learners at respective points in time. It includes tracking of the learners' progress and other future learning effects, as well as the dropouts. It creates a new benchmark for stakeholders to choose the best practices other than mere learning process; where ever organizations are held responsible for the attritions, satisfaction, or failure of these learners.

C. Enhanced market analysis: This shows that organizations can complement the market analysis using educational big data by focusing on rates of success, accomplishments, and areas of deficiency with corresponding indexed institutions. This data is also useful in the assessment of business performance as well as key performance indicators and academic accreditations. It also assists in explaining the achievement, and development opportunities, and to influence strategic and innovative approaches. Besides, organizations can assess their learners, teachers, and curriculum outputs against comparable organizations for better knowledge of what to expect in the next improvements and organizational strategies.

E. Predictive teaching and assessment strategy: The application of educational big data enables teachers to get prompt and accurate assessments of their content delivery and strategies of pedagogy and evaluation. Therefore, the observation of learners' capabilities as well as their level of knowledge enables the teachers to follow the learning process to identify vulnerability and likely failure areas. Educational big data also enables the identification of checkpoints/defects in educational programs/courses or any type of content for using big data analysis and understanding the depth of improving the curriculum. It supports differentiated instruction, immediate check and feedback, learning activities, and group learning. Besides, it evaluates learners' performance based on previous records which makes, allowing it to forecast their success.

4 Educational Data Mining

Data Mining is a powerful tool that can discover valuable information through analyzing data from different perspectives or views, putting that data into understandable forms and distilling the intrinsic pattern embedded in the data base [9]. It thereby enables enhanced decision-making to occur in the long run. Data mining involves the processes of descriptions, classification, clustering, forecasting, association, and others.

4.1 Goals of Educational Data Mining

Below are the main targets set for using EDM [10]:

A. Predict Student Behavior: This goal can be attained through the creation of models of students, which describe and classify the attributes or conditions of a student, including, awareness, attitudes, learning motivation, knowledge, and meta-cognition characteristics.

B. Enhance Knowledge Models: Enhance understanding and categorization of education concepts through employ of EDM techniques. EDM assists in building improved models of knowledge through examination of patterns in the data provided by students, about their engagement with aspects such as learning material, and students' performances in various subjects. These models help educators to know how students understand various facts, which connections between facts are essential, and how knowledge should be arranged to facilitate understanding and assimilation of knowledge.

C. Evaluate Learning Support Systems: Review several different educational tools in order to identify and assess the efficacy of the methods that will enhance the student's performance.

D. Advance Educational Research: Through EDM, one can discover the learning profile of the student, aspects of learning that affect success or failure, and which instructional approaches are effective. This data analysis assists in identifying new patterns in students' behavior and learning processes thus leading to changes in educational practices. Consequently, it ensures the progress of educational theories and increases the reliability of educational systems and methods based on empirical data.

4.2 Educational Data Mining Settings

Over recent years, EDM has gained significant attention from researchers across various global disciplines, involving different groups of users or participants.

The data is collected from different distributed sources; it can be structured and unstructured data types. This data primarily originates from two sources [10] [11]: offline and online. Offline data comes from traditional and modern classrooms, interactive teaching and learning environments, learner information, student attendance, emotional data, course details, and academic records from institutions. Online data, on the other hand, is derived from geographically dispersed educational participation, distance education, web-based learning, and computer-assisted collaborative learning through social networking sites and online forums. The application of data mining in an educational system is illustrated in Figure

3 [12]. It is generally referred to as knowledge discovery in databases. In this process, knowledge is extracted or mined from vast amounts of educational data, including data from teachers, resources, students, alumni, etc. Academics and educators use EDM to design, plan, build, and maintain educational systems to improve student performance. Students interact directly with these systems, generating data that serves as input for EDM. The system provides students with recommendations and uncovers new insights for educators by employing a range of data extraction methods, including clustering, classification, and pattern comparisons.

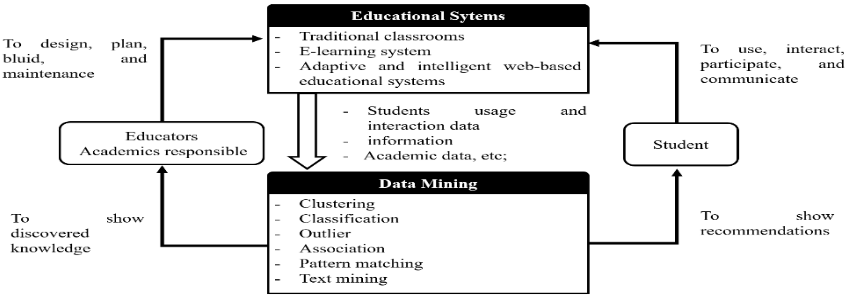


FIG 3: The cycle of data mining in the educational system

4.3 Educational Data Mining Settings

EDM employs a specific set of data mining methods according to the application and objective for which they are used in the EDM process (Table 2). These methods can be classified as shown (Figure 4) [13].

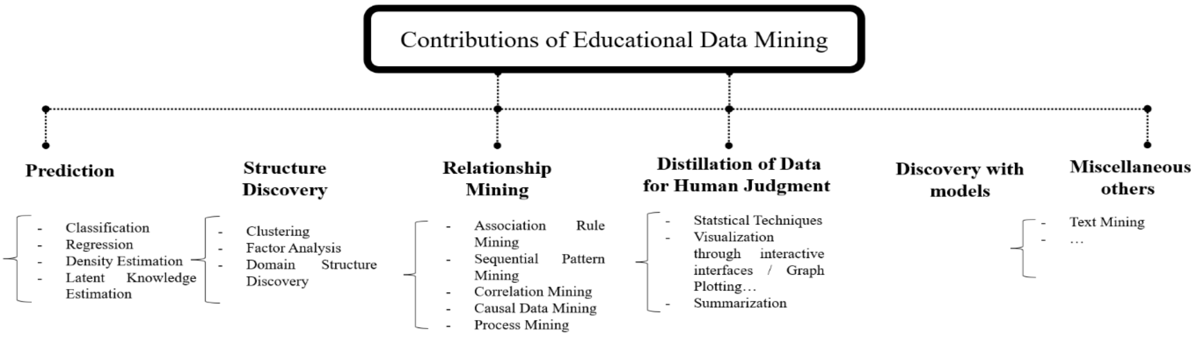
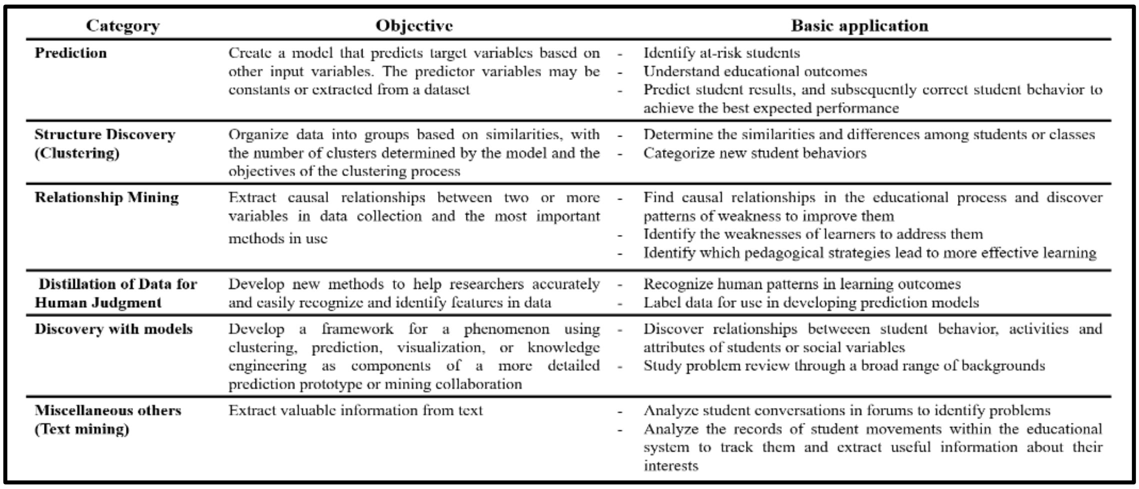


FIG 4: Classification of EDM methods



Category	Objective	Basic application
Prediction	Create a model that predicts target variables based on other input variables. The predictor variables may be constants or extracted from a dataset	- Identify at-risk students - Understand educational outcomes - Predict student results, and subsequently correct student behavior to achieve the best expected performance
Structure Discovery (Clustering)	Organize data into groups based on similarities, with the number of clusters determined by the model and the objectives of the clustering process	- Determine the similarities and differences among students or classes - Categorize new student behaviors
Relationship Mining	Extract causal relationships between two or more variables in data collection and the most important methods in use	- Find causal relationships in the educational process and discover patterns of weakness to improve them - Identify the weaknesses of learners to address them - Identify which pedagogical strategies lead to more effective learning
Distillation of Data for Human Judgment	Develop new methods to help researchers accurately and easily recognize and identify features in data	- Recognize human patterns in learning outcomes - Label data for use in developing prediction models
Discovery with models	Develop a framework for a phenomenon using clustering, prediction, visualization, or knowledge engineering as components of a more detailed prediction prototype or mining collaboration	- Discover relationships betweeen student behavior, activities and attributes of students or social variables - Study problem review through a broad range of backgrounds
Miscellaneous others (Text mining)	Extract valuable information from text	- Analyze student conversations in forums to identify problems - Analyze the records of student movements within the educational system to track them and extract useful information about their interests

TAB 2: Objectives and applications of EDM methods

4.4 Educational Data Mining Settings

The figure 5 illustrates the various ways in which EDM contributes to the field of education, emphasizing its ability to enhance learning processes. EDM promotes collaborative learning, predictive analysis, behaviour modeling, and data visualization, providing a thorough structure for evaluating and enhancing student achievement, engagement, and retention. Through the use of sophisticated data analysis, EDM supports well-informed decision-making for educators, enabling the creation of customized learning methods and the enhancement of teaching approaches. By incorporating data-driven insights into education, EDM demonstrates its capacity to revolutionize traditional educational practices and improve overall learning results [14].

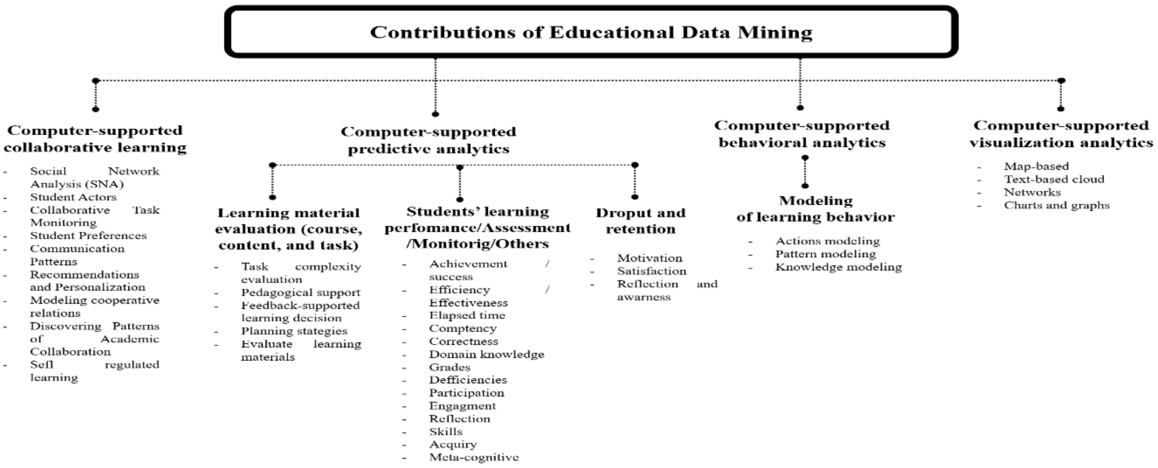


FIG 5: Educational Data Mining contributions

A. Computer-supported collaborative learning (CSCL): The integration of EDM into CSCL conclusively improves the efficacy of collaborative learning through the use of data mining technologies to analyze grouping interactions in Learning Management Systems (LMS) among learners. This enables educators to evaluate group interaction, monitor behaviors and attend to any interventions in sequence. Based on the data collected from discussion forums, content delivery, and assessments, instructors can identify students' preferences and recognize if they are out of control in learning process [15]. EDM thus enables interventions to be provided as and when needed, ensuring that cooperation between stakeholders as well as the students is enhanced as well as more desired learning outcomes are achieved. EDM delivers students' learning patterns and models that can be followed in order to develop individual learning-teaching schemes to foster cooperation and students' cognition. EDM not only contributes to the self-identity and interactivity but also contributes to the better development of the effectiveness of the collaborative learning environment in creating a better CSCL result for learners and their groups needed by society:

- **Social Network Analysis (SNA):** is a helpful technique in EDM that displays the interactions of students and their connections. In CSCL, SNA can explain the structure of group collaboration, who are the core members, commonly referred to as triad active students, and who is isolated or considered an isolate. With this information on hand, a teacher can better comprehend the general functioning of a group and intervene where necessary [16] [17].
- **Student Actors:** As students are seen as 'actors' in the collaborative tasks, EDM tracks the student's contribution, participation, and efficiency in the group tasks. Information gathered in regard to student behavior may reflect certain aspects such as communication, problem-solving, or leadership among

others. For instance, students whose participation is always in leadership positions may be motivated to offer mentorship to other learners while others may require the prompts to be more active [18].

- **Collaborative Task Monitoring:** EDM examines the performance of students in collaborative tasks using data of the different groups of students. For example, it can determine which collaborative tasks foster deep learning versus those that lead to superficial engagement. When the effective collaboration patterns are known teachers can adjust the structure of subsequent tasks in a manner applied to foster effective interaction among the group members [19].
- **Student Preferences:** EDM can gather relevant information to recognize and interpret student's attitudes and interest concerning different aspects of collaboration such as preferences for synchronous versus asynchronous communication, task allocation, or leadership roles. This makes it possible to undertake personalized adjustments in CSCL, to meet the needs of each student since the system is knowledgeable about students' learning preferences leading to formation of proper tasks or grouping patterns [20].
- **Communication Patterns:** EDM is used for monitoring participation levels in group interactions, detecting communication bottlenecks and analyzing level of dialogue quality to ensure equitable and meaningful contributions within CSCL environment. It provides tips on joint work, promotes peer learning and supports studying how students communicate their data that the style is improved to efficiently allocate tasks. In addition, EDM can provide instantaneous and real-time feedback on what the group attention should be focused upon — engagement (or lack thereof) or topic importance. Ultimately, helping to provide greater cognitive and social engagement to deliver more effective experiences in collaborative work [21] [22].
- **Recommendations and Personalization:** Data analysis through EDM enables tracking of individual and group interactions, learning styles and performance metrics which offer real-time recommendations to both students and educators for a better collaborative experience. That is, EDM can recommend individualized learning content, groupings of students based on their skill profiles, and manipulative experience using complementary strengths among them or tasks that cater to the strength profile of each student. It also helps to know who the underperforming students or inactive members are and provides them with personalized interventions to get back on track. EDM uses data to personalize the learning journey while providing actionable recommendations to support students in doing well-interleaved and group-based tasks [23] [24].
- **Modeling cooperative relations:** EDM supports CSCL through modeling and simulating collaborative relations between students. EDM can model student collaboration, roles, dependencies, and strength of cooperative ties by tracking interactions, task contributions and communication patterns. It looks at patterns of trust, reciprocity, and dependence between students to predict success or potential failure in collaboration. These lead to better performance, in terms of both efficiency (rate of user completion) and equality (variance between users' scores). Consequently, educators can make optimal groups at the outset, spread tasks more evenly among learners beforehand and interfere when necessary to guarantee effective collaboration [25].
- **Discovering Patterns of Academic Collaboration:** EDM can find academic collaboration patterns that lead to effective learning outcomes by mining data from collaborative activities. Frequent collaboration amongst high-achieving students, for instance, may lead to improved group performance; however, imbalanced collaboration may indicate problems such as "free-riding," in which one or two students do all the work. Understanding these patterns helps in designing better collaborative tasks and intervention strategies [26].
- **Sefl regulated learning:** EDM uses learning analytics dashboards and provides actionable insights for promoting self-regulated learning by unleashing the potential of students about the information related to engagement, progress, and performance. It allows the students to keep a check on their study

time and set effective goals. Similarly, EDM also helps to set goals and manage time, by visualizing how students consume their study time and where potential changes can be made. EDM internally uses predictive analytics to predict potential problems, which will only benefit students pre-emptively. In addition, insights on engagement illustrate trends where students may be disengaging or slowing down their progress, so that students can self-correct and return to course learning objectives [27] [28] [29].

B. Computer-supported predictive analytics (CSPA):

- **Learning Material Evaluation:** EDM is invaluable in informing feedback-oriented decision-making to support enhanced quality in course, content and tasks in terms of learning material evaluation. By analyzing performance and interaction data that students have with learning materials, EDM can provide intelligence on how well students are engaging with learning resources, which will serve as intelligence of task complexity (i.e., identifying which tasks are appropriately complex and which ones can be adjusted). This is critical for educators to identify planning of how to organize and present learning content for effective learning. Beyond creating engaging learning content, EDM can also identify areas of student difficulty, allowing educators to establish targeted (i.e., pedagogical) learning supports. This is an ongoing process that facilitates improvements of the learning materials to accommodate student learning needs in a dynamic environment that will enhance educational outcomes [30] [31].
- **Students' Learning Performance:** EDM offers an enriching analysis of student learning performance by monitoring key measures through the indicators of time spent on learning tasks, proficiency taken on and accuracy of response, and student engagement across their learning opportunity as a whole. For example, through EDM, achievement and engagement can be monitored to provide an estimate of how students are developing and any strengths and weaknesses in the learning experience. Furthermore, it can assess learning efficiency by considering speed and correctness in the completion of tasks, and also explore meta-cognitive elements such as reasoning and inquiries. Inception, EDM captures GPA, skills growth, and other indicators to give a comprehensive account of a student’s learning experience thus enhancing the provision of authorized interventions and personalized attention to foster improved learning performances alongside academic outcomes [32] [33].
- **Dropout and Retention:** EDM provides insights that explain drivers related to dropout and retention, interpreting important facets such as motivation, satisfaction, self-reflection, and awareness. With the examination of student data, EDM establishes a sense of student motivation, satisfaction with the environment, and self-reflection of academic progress. Furthermore, assessments extend to students' awareness of their performance and factors that need to be identified for improvement. With an understanding of all the contributors, educators can identify early warning signs of a possible dropout and initiate a more effective intervention. The intervention can personalize and support student needs to increase retention and assist in a more positive educational experience overall [34] [35].

C. Computer-Supported Behavioral Analytics (CSBA): With data mining techniques can discover significant findings and objects in learning patterns from how diverse students learn. Several research studies have illustrated the benefits of utilizing online learning analytics and particularly in collaborative remote environments to improve students' experiences. Instead, we are currently focusing on deploying real-time data to support agencies to monitor and improve the learning process in order for the students to solve problems with various complexities. We already have models that predict learning processes, taking into account the knowledge of students, motivation and attitudes. Interactions and activities in online environments can also hint at emerging learning behaviors, or predict outcomes (and therefore needs for support) [36] [37].

D. Computer-Supported Visualization Analytics (CSVA): CSVA integrates information visualization with advanced data mining and knowledge representation to permit visual analysis of individual behaviour associated with activities. In the educational domain, CSVA takes advantage of visual data mining tools to provide analytic views about learning processes and the learner experience. To illustrate, it visually represents online discussions, and then measures the quality and value of the posts using systematic algorithms to identify postings with instructional value. Additionally, visual data mining tools could provide more flexibility and diversity in the assessment of higher education environments contributing to beneficial learning outcomes. Finally, visual representations of students' engagement enable instructors to evaluate students' online behaviour’s, as well as organize and track data in online learning environments. Visual data mining tools also assist with complex data representation, and can easily track students' activity in online learning environments [38] [39].

4.5 Educational Data Mining Contributions Across Methods Used

The table below provides a detailed overview of different data mining techniques across various dimensions of EDM applications. The table categorizes these applications into four broad areas: Computer-Supported Collaborative Learning, Computer-Supported Predictive Analytics, Computer-Supported Behavioral Analytics, and Computer-Supported Visualization Analytics. The alignment facilitates the process of choosing the right methods to extract insights about learning processes and prediction results (Table 3) [14].

Based on the literature review conducted and the results from Table 3, the figure 6 represents the most commonly used methods in EDM.

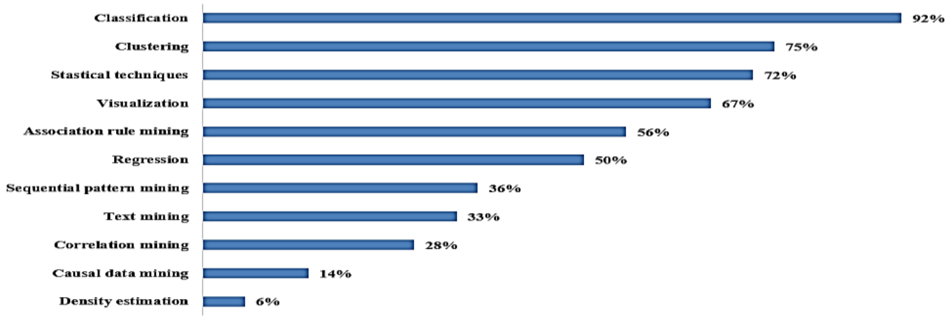


FIG 6: Most commonly used EDM Methods

Method / Aspect	Prediction			Structure Dicovery	Relationship Mining				Human Judgment		Miscellane-ous others
	Classification	Regression	Density estimation	Clustering	Association rule mining	Sequential pattern mining	Correlation mining	Causal data mining	Statstical Techniques	Visualization	Text mining
Computer-supported collaborative learning											
Social Network Analysis						✓					
Student Actors	✓			✓					✓		
Collaborative Task Monitoring	✓			✓			✓		✓	✓	✓
Student Preferences	✓	✓		✓		✓	✓		✓		
Communication Patterns	✓								✓	✓	✓
Recommendations and Personalization	✓			✓	✓	✓			✓		
Modeling cooperative relations				✓		✓				✓	
Discovering Patterns of Academic Collaboration	✓				✓	✓			✓	✓	✓
Sefl regulated learning	✓				✓				✓	✓	✓
Computer-supported predictive analytics											
Learning material evaluation (course, content, and task)											
Task complexity evaluation	✓	✓		✓	✓						
Pedagogical support	✓			✓	✓					✓	✓
Feedback-supported learning decision	✓			✓	✓		✓		✓	✓	✓
Planning strategies	✓			✓	✓	✓				✓	
Evaluate learning materials	✓	✓	✓	✓	✓				✓	✓	✓
Droput and retention											
Motivation	✓			✓	✓				✓		
Satisfaction	✓	✓		✓					✓		
Reflection and awarness	✓			✓					✓	✓	

Aspect \ Method	Prediction			Structure Dicovery	Relationship Mining				Relationship Mining		Miscellane-ous others
	Classification	Regression	Density estimation	Clustering	Association rule mining	Sequential pattern mining	Correlation mining	Causal data mining	Statstical Techniques	Visualization	Text mining
Students' learning perfomance/Assessment/Monitorig/Others											
Achievement / success	✓	✓		✓	✓	✓	✓	✓	✓	✓	
Efficiency / Effectiveness	✓	✓		✓					✓	✓	
Elapsed time	✓	✓									
Comptency				✓					✓		
Correctness	✓	✓		✓	✓		✓		✓		
Domain knowledge	✓	✓		✓	✓			✓	✓	✓	✓
Grades	✓	✓		✓	✓		✓		✓	✓	
Defficiencies	✓	✓			✓					✓	
Participation	✓	✓							✓	✓	✓
Engagment	✓	✓		✓					✓	✓	
Reflection	✓					✓			✓	✓	✓
Skills	✓	✓			✓				✓	✓	
Acquisition	✓	✓		✓	✓	✓				✓	
Acquiry	✓			✓							✓
Meta-cognitive	✓			✓		✓				✓	
Computer-supported behavioral analytics											
Modeling of learning behavior											
Actions modeling	✓	✓		✓	✓	✓	✓	✓	✓	✓	
Pattern modeling	✓	✓		✓	✓	✓	✓		✓	✓	
Knowledge modeling	✓	✓		✓	✓	✓	✓	✓	✓		
Computer-supported visualization analytics											
Map-based, text-based cloud, networks and charts and graphs	✓		✓	✓	✓		✓	✓	✓	✓	✓

TAB 3: EDM contributions across various data mining methods use

5 Educational Big Data & Educational Data Mining Challenges

In addition to the prospective benefits of big data and EDM, there are drawbacks and limitations [12] [40] [41].

- **Data Quality Control:** Educational data might be incorrect, incomplete, inaccurate, or even completely missing, which complicates the analysis.
- **Privacy and Security:** EDM and educational big data require the management of personally sensitive data about students and educators thereby raising privacy and ethical practice questions about using this data. There is a need to adequately ensure that collecting, storing, and using educational data provides privacy and security for individuals.
- **Storage:** Educational datasets may become huge in size or volume to the degree that required resource requirements storage make it challenging to handle educational data.
- **Formatting and Data Cleaning:** are basic and fundamental data management processes and yet require a certain threshold of technical skills depending upon the volume and complexity of the big data. Formatting and data cleaning are practices needed to eliminate idiosyncrasies reduce errors, and set up the data in a manner to obtain important analysis and information gain toward informed decision-making.
- **Platform Integration:** presents its own set of challenges, ranging from compatibility of data formats and structures across platforms thereby creating data duplicates and discrepancies concerning data integration. These issues result in duplicate activities, additional high costs, and wasted resources when departments struggle to coordinate and or standardize their data sources/feeds more effectively.
- **Technical expertise:** A limitation of EDM is that it requires complex technical expertise. It employs complex and advanced sophisticated systems-based data analysis and artificial intelligence techniques that are often out of reach for the common, non-technical educational worker.
- **Interpreting the findings to make sense of them:** A further challenge EDM faces is that the results of the analyses typically are complicated and often difficult to understand. The results can be convoluted which also makes it difficult understand what the clear purpose and orientation of the results are.

6 Conclusion

The combination of EDM and educational big data within the process of teaching and learning provides tremendous opportunities to change education. As traditional models of teaching become less fit for purpose, EDM offers us a means to create more efficient and flexible employment of educational practices. EDM could increase the understanding of student behavior and the learning environment when using considerable amounts of educational data, which can foster data-driven decisions that improve educational outcomes. This review suggested that EDM has the potential to innovate practice in education and has current applications in predicting academic performance, recommending individualized pathways of learning, identifying undesirable behaviors, and developing student profiles. The implications of this research are clear, we should adopt advanced data analytics to adapt teaching practice to the new environment in education.

In keeping with the goal of innovating practice through educational change we propose to develop a teaching model that situates EDM techniques with Lean Management tools to further enhance educational practices. By adopting the quality of education flow and efficiency, we may improve educational learning quality and subsequently align teaching practices with student and industry demands. This teaching model aims to form the basis of a responsive and reflective educational environment that consistently seeks to obtain the best learning outcomes.

7 References

[1] Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020b). Big data in education : a state of the art, limitations, and future research directions. International Journal Of Educational Technology In Higher Education, 17(1).

[2] Lesjak, D., & Kohun, F. (2021b). BIG DATA ANALYTICS IN HIGHER EDUCATION. Issues In Information Systems.

[3] Barbeiro, L., Gomes, A., Correia, F. B., & Bernardino, J. (2024b). A Review of Educational Data Mining Trends. Procedia Computer Science, 237, 88 95.

[4] Kitchenham, B., & Brereton, P. (2013b). A systematic review of systematic review process research in software engineering. Information And Software Technology, 55(12), 2049 2075.

[5] Bai, X., Zhang, F., Li, J., Guo, T., Aziz, A., Jin, A., & Xia, F. (2021). Educational Big Data : Predictions, Applications and Challenges. Big Data Research, 26, 100270.

[6] Zhang, X., Sun, G., Pan, Y., Sun, H., He, Y., & Tan, J. (2018). Students performance modeling based on behavior pattern. Journal Of Ambient Intelligence And Humanized Computing, 9(5), 1659 1670.

[7] Raza, N. M., RKayani, N. H. U., Malik, N. A., Gul, N. M., & Suleman, N. A. (2023). TRENDS AND APPLICATIONS OF BIG DATA IN EDUCATION. Pakistan Journal of Science, 75(02), 345 352.

[8] Hazra, S., & Ganguli, S. (2021b). Big Data and Its Application in Smart Education during the COVID-19 Pandemic Situation. Dans CRC Press eBooks (p. 187 201).

[9] Fuseini, I., & Missah, Y. M. (2024). A critical review of data mining in education on the levels and aspects of education. Quality Education for All, 1(2), 41 59.

[10] Bachhal, P., Ahuja, S., & Gargrish, S. (2021). Educational Data Mining : a review. Journal Of Physics Conference Series, 1950(1), 012022.

[11] Mehra, C., & Agrawal, R. (2020). Educational data mining approaches, challenges and goals : A review. JIMS8I - International Journal Of Information Communication And Computing Technology, 8(2), 442 447.

[12] Haryani, C.A., Daicy, W., Widjaja, A.E., Aribowo, A. and Hery, K. (2022), “Educational data mining: the application in the university’s feedback survey analysis using classification and clustering techniques”, 2022 International Conference on Science and Technology, ICOSTECH 2022,

[13] Aleem A, Gore MM. Educational data mining methods: A survey. In: 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT). Gwalior, India: IEEE; 2020. pp. 182-188

[14] Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education : A review and synthesis. Telematics And Informatics, 37, 13 49.

[15] Garshasbi, S., Yecies, B., & Shen, J. (2021). Microlearning and computer-supported collaborative learning : An agenda towards a comprehensive online learning system. STEM Education, 1(4), 225.

[16] Benabdelouahab, S., Bouhdidi, J. E., Younoussi, Y. E., & De Gea, J. M. C. (2022). Learning Styles Prediction Using Social Network Analysis and Data Mining Algorithms. Dans Lecture notes on data engineering and communications technologies (p. 315 322).

[17] Samanta, P., Sarkar, D., Kole, D. K., & Jana, P. (2020). Social Network Analysis in Education : A Study. Dans Intelligent systems reference library (p. 65 83).

[18] Jie, W., Hai-Yan, L., Biao, C., et al.: Application of educational data mining on analysis of students' online learning behavior. In: 2017 2nd International Conference on Image, Vision and Computing (ICIVC). IEEE, (2017)

[19] AlQaheri, H., & Panda, M. (2022). An Education Process Mining Framework : Unveiling Meaningful Information for Understanding Students' Learning Behavior and Improving Teaching Quality. Information, 13(1), 29.

[20] S. Alturki and N. Alturki, "Using Educational Data Mining to Predict Students' Academic Performance for Applying Early Interventions," JITE:IIP, vol. 20, pp. 121–137, 2021

[21] Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., & Gašević, D. (2019). Orchestrating Learning Analytics (ORLA) : Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. Australasian Journal Of Educational Technology, 35(4).

[22] Romero, C., & Ventura, S. (2020b). Educational data mining and learning analytics : An updated survey. Wiley Interdisciplinary Reviews Data Mining And Knowledge Discovery, 10(3).

[23] Xiong, Z., Li, H., Liu, Z., Chen, Z., Zhou, H., Rong, W., & Ouyang, Y. (2024b). A Review of Data Mining in Personalized Education : Current Trends and Future Prospects. ArXiv (Cornell University).

[24] Fang, C., & Lu, Q. (2021). Personalized Recommendation Model of High-Quality Education Resources for College Students Based on Data Mining. Complexity, 2021, 1 11.

[25] Hidayat, N., Wardoyo, R., & Azhari, S. (2018). Educational Data Mining (EDM) as a Model for Students' Evaluation in Learning Environment. 2018 Third International Conference On Informatics And Computing (ICIC).

[26] Hung, H.C., Liu, I.F., Liang, C.T. and Su, Y.S., 2020. Applying educational data mining to explore students' learning patterns in the flipped learning approach for coding education. Symmetry, 12(2), p.213

[27] Nuankaew, P. (2022). Self-Regulated Learning Model in Educational Data Mining. International Journal Of Emerging Technologies In Learning (iJET), 17(17), 4 27.

[28] Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & Education, 104, 18 33.

[29] Kim, K., & Moon, N. (2018). Activity index model for self-regulated learning with learning analysis in a TEL environment. The Journal Of Supercomputing, 75(4), 1971 1989.

[30] Rabelo, A., Rodrigues, M. W., Nobre, C., Isotani, S., & Zárate, L. (2023). Educational data mining and learning analytics : a review of educational management in e-learning. Information Discovery And Delivery, 52(2), 149 163.

[31] Rodrigues, M. W., Isotani, S., & Zárate, L. E. (2018). Educational Data Mining : A review of evaluation process in the e-learning. Telematics And Informatics, 35(6), 1701 1717.

[32] Yağcı, M. (2022). Educational data mining : prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1).

[33] Chaka, C. (2022). Educational data mining, student academic performance prediction, prediction methods, algorithms and tools : an overview of reviews. DOAJ (DOAJ : Directory Of Open Access Journals).

[34] Triayudi, A. (2023). Students' drop out identification using educational data mining. AIP Conference Proceedings.

[35] Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student Retention Using Educational Data Mining and Predictive Analytics : A Systematic Literature Review. IEEE Access, 10, 72480 72503.

[36] Feng, G., & Fan, M. (2024). Research on learning behavior patterns from the perspective of educational data mining : Evaluation, prediction and visualization. Expert Systems With Applications, 237, 121555.

[37] Mourabit, I. E., Jai-Andaloussi, S., & Abghour, N. (2021b). Educational Data Mining Techniques for Detecting Undesirable Students' Behaviors and Predicting Students' Performance : A Comparative Study. Dans Advances in intelligent systems and computing (p. 163 170).

[38] V. Siafis and M. Rangoussi, ''Educational data mining-based visualization and early prediction of student performance: A synergistic approach,'' in Proc. 26th Pan-Hellenic Conf. Informat., Nov. 2022, pp. 246–253.

[39] Sharma, P., & Sharma, S. (2020b). DATA MINING TECHNIQUES FOR EDUCATIONAL DATA : A REVIEW. International Journal Of Engineering Technologies And Management Research, 5(2), 166 177.

[40] Debang, M., & Hassan, B. U. (2023d). Educational Data Mining : Prospects and Applications. Authorea (Authorea).

[41] Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining Big Data in Education : Affordances and Challenges. Review Of Research In Education, 44(1), 130-160. https://doi.org/10.3102/0091732x20903304.

Artificial Intelligence Tools and Bin Packing Problem: A Literature Review

Nohaila Belhaj[1], Abdelfettah Sedqui[1], Najlae AlFathi[1]

[1] Laboratory of Innovative Technologies National school of applied sciences, Abdelmalek Essaadi University, Tangier, Morocco

nohaila.belhaj555@gmail.com

abdelfettah.sedqui@gmail.com

najlae.alfathi@gmail.com

Abstract

Bin packing problem has widespread applications in industries and logistics and has become the focus of attention of researchers. In literature, there exist exact, heuristic and metaheuristic methods dedicated to solving this problem, but an exact optimal solution doesn't exist, especially for large instances. However, the utilization of the new technologies related to artificial intelligence (AI), such as machine learning (ML) and Deep learning (DL), enhances the solutions in terms of quality and time optimization. In this paper, we investigate how artificial intelligence (AI) can make this improvement. This review followed the Preferred Reporting Items for Systematic Reviews (PRISMA) standards to extract peer-reviewed papers from research databases between 2020 and 2024. The analysis considered 27 articles judged relevant to our topic. The results showed that the application of artificial intelligence (AI) tools to solve bin packing problems (1D, 2D and 3D) gives robust results compared to other methods.

Keywords: *Bin packing, Artificial intelligence, machine learning, deep learning, improvement*

1 Introduction

Starting from Covid-19 pandemic effects in boosting the e-commerce [1] in recent years and the increase in demand, the logistics is becoming a crucial aspect in our life, especially, the packing of different shapes and weight of items into containers. Scientifically, bin packing problem transforms the packing issues to mathematical formulation including all constraints related to weight, dimensions, capacity, etc. and consists of packing all items into minimum number of containers orthogonally [2].

From decades, the researchers proposed a lot of methods to solve 1D, 2D and 3D bin packing problems, which could be classified into three main categories:

Exact algorithms and lower bounds: several techniques can be found in the works of Pisinger and Sigurd [3], Puchinger & Raidl [4], Sgall [5], Martello & Vigo [6], Fekete & Schepers [7] and [8].

Heuristics Algorithms: either one phase algorithms such as the ones presented by Berkey & Wang [9], Bhatia & al. [10] or two-phase algorithms proposed by Chung, & al. [11], Lodi & al. [12], Monaci & Toth [13] and Dai & J. Cha [14].

Metaheuristics Algorithms: Tabu search, Ant Colony Optimization, Genetic Algorithm and other algorithms present metaheuristics methods that can be applied to solve different types of Bin Packing problem and to generate optimal solutions by exploring a large search space efficiently. The reader can refer to [15], [16], [17], [18], [19], and [20] for more details.

On the other hand, the apparition of AI and ML make a revolution in all domains, finance [21], healthcare [22], education [23], industry 4.0 [24], logistics and supply chain [25], including the optimization of bin packing problems.

In the same context, our study focuses on the bin packing problem solutions improvement based on AI and ML techniques, thus, the answer is provided from the analysis of the current scientific databases (Scopus, ScienceDirect, IEEE and Springer link) peer-review articles published between 2020-2024.

The remainder of the paper is organized as follows: Section 2 summarizes the bin packing problem definition, and AI evolution, Section 3 describes the review methodology and results, followed by a discussion in Section 4, then, conclusion and future perspectives is presented in section 5.

2. Bin packing problem and artificial intelligence (AI) evolution

2.1 Bin packing problem

Bin packing problem is a classic optimization problem which aim to pack a set of items with various size into minimum number of bins with fixed capacity. Generally, the problem can be treated trough (1D, 2D, and 3D):

2.1.1 One dimensional bin packing

This version is a NP-hard optimization problem [26], where the problem will optimize the length or the width of the bin in a direction [27] which can be formulated as follow [28]:

Consider a set of n items with size w_j to pack and a set of u bins available with a fixed capacity C, two binary decision variables are presented as:

$$y_i = \begin{cases} 1 & if\ bin\ i\ is\ used\ in\ the\ solution \\ 0 & otherwise \end{cases}$$

$$For \quad i \backslash in \{1, \ldots, u\}$$

$$x_{ij} = \begin{cases} 1 & if\ item\ j\ is\ packed\ into\ bin\ i \\ 0 & otherwise \end{cases}$$

$$For \quad i \backslash in \{1, \ldots, u\}\ and\ j \backslash in \{1, \ldots, n\}$$

Minimize $$\sum_{i=1}^{u} y_i \tag{1.1}$$

s.t. $$\sum_{j=1}^{n} w_j \,.\, x_{ij} \ \leq C . y_i \qquad For \quad i \backslash in \{1, \ldots, u\} \tag{1.2}$$

$$\sum_{i=1}^{u} x_{ij} = 1 \qquad For \quad j \backslash in \{1, \ldots, n\} \tag{1.3}$$

$$y_i \in \{0, 1\} \qquad For \quad i \backslash in \{1, \ldots, u\}$$

$$x_{ij} \in \{0, 1\} \qquad For \quad i \backslash in \{1, \ldots, u\}\ and\ j \backslash in \{1, \ldots, n\}$$

Equation (1.2) imposes the non-exceeding of the bin's capacity, and each item must be packed in one and only one bin (1.3), while in the equation (1.1) the objective function is described aiming to minimize the number of bins used.

2.1.2 Two-dimensional bin packing

This problem is a strongly NP-hard optimization problem since it is a generalization of the one-dimensional bin packing problem and it consists of two dimensions instead of one (e.g. width and height (W, H)), and it is formulated by considering an enumerated set of items called "patterns", represented by a binary vector Vj = (vij), to be packed into a single bin [29]:

Consider a set of n items, m subsets of items (patterns) and H as the height of the bin and Vj = (vij) is a binary decision vector defined by:

$$v_{ij} = \begin{cases} 1 & if\ the\ item\ i\ belongs\ to\ the\ pattern\ j \\ 0 & otherwise \end{cases}$$

$$For \quad i \backslash in \{1, \ldots, n\}\ and\ j \backslash in \{1, \ldots, m\}$$

Minimize $$\sum_{j=1}^{m} x_j \tag{2.1}$$

s.t. $$\sum_{j=1}^{m} v_{ij} \,.\, x_j \ \leq 1 \qquad For \quad i \backslash in \{1, \ldots, n\} \tag{2.2}$$

$$\sum_{i=1}^{n} v_{ij} \,.\, h_j \ \leq H \qquad For \quad j \backslash in \{1, \ldots, m\} \tag{2.3}$$

$$x_j \in \{0, 1\} \qquad For \quad j \backslash in \{1, \ldots, n\} \tag{2.4}$$

The objective function presented in equation (2.1) sights to minimize the number of created patterns, thus, in the equation (2.2) each item must belong to only one pattern, then, in the equation (2.3) each pattern must not exceed the second dimension (the height H) of each bin, and the equation (5.5) is presenting a binary variable decision defined as:

$$x_j = \begin{cases} 1 & \textit{if pattren j is a solution to the problem} \\ 0 & \textit{otherwise} \end{cases} \quad ; \quad \text{For} \quad j \mid in \{1, \ldots, n\}$$

2.1.3 Three-dimensional bin packing

This version is also a strongly NP-hard optimization problem which consists to pack a set of rectangular cuboidal items within the minimal number of bins without exceeding their dimensions [30] and it can be formulated as [31]:

Consider a set of *n* items with dimensions l_i, w_i, h_i (resp. length, width, height) and a volume $V_i = l_i * w_i * h_i$ to pack, and a bin with a dimensions L, W, H (resp. length, width, height) and a volume $V_c = L * W * H$, and two decision variables are presented as:

x_i, y_i, z_i : coordinates of the bottom-left corner of item *i* within the bin

$$\lambda_i = \begin{cases} 1 & \textit{if item i is packed within the box} \\ 0 & \textit{otherwise} \end{cases}$$

$$\text{Maximize} \quad \frac{\sum_{i=1}^{n} V_i \cdot \lambda_i}{V_c} \tag{3.1}$$

$$\text{s.t.} \quad x_i + l_i.\lambda_i \le x_j \quad \text{or} \quad y_i + w_i.\lambda_i \le y_j \quad \text{or} \quad z_i + h_i.\lambda_i \le z_j$$

$$\textit{For} \quad i,j \mid in \{1, \ldots, n\} \quad (i \neq j) \tag{3.2}$$

$$0 \le x_i \le L - l_i.\lambda_i \quad \text{or} \quad 0 \le y_i \le W - w_i.\lambda_i \quad \text{or} \quad 0 \le z_i \le Z - h_i.\lambda_i \tag{3.3}$$

$$\sum_{i=1}^{n} V_i . \lambda_i \le V_c \quad ; \quad \textit{For} \quad i \mid in \{1, \ldots, n\} \tag{3.4}$$

The equation (3.1) presents the objective function that aims to maximize the utilization of bin's volume, thus, the equation (3.2) sights to avoid overlapping between the packed items, then, the equation (3.3) ensures that the whole item is contained within the dimensions of the bin, and the equation (3.4) guarantees that the volume of the bin is not exceeded.

2.2 AI evolution

The concept of AI was early developed by Alan Turing and John von Neumann, and formally introduced by John McCarthy at the Dartmouth Conference in 1956 [32]. The 1990s experienced the "intelligent agent" inspired from decision-making abilities of human, while the reduced uncertainty in decision-making was achieved in the late of 1990s with the utilization of machine learning techniques [33]. The 21st century has perfectly boomed AI into an incredible revolution, reinforced by ML, DL, and artificial neural networks (ANN).

In literature several works were published presented new methods-based AI and ML to solve bin packing problem, proving their efficiency for this kind of problem solving.

In 1990s, two categories of new techniques were proposed:

Single-solution methods: that consist of improving one candidate solution such as Simulated annealing (SA) proposed by Kirkpatrick & al. [34], inspired from the concept of fluids solidification process and tested on 1D bin packing problem by Drexl [35].

Population based methods: which improve a set of solutions through several operations like the genetic algorithm (GA) introduced by Holland [36], inspired from the natural evolution theory and improve the solutions via several generations according to the natural process including selection, crossover and mutation, and applied by Khuri & al. [37].

In the 2000s, other methods were born, principally, Ant colony optimization (ACO) which follows the ant's behavior on searching foods through the shortest path [38], particle swarm optimization (PSO) inspired from particles' (birds, fish, …) behavior and movement [39], and especially the hybrid AI algorithm which improves perfectly the solutions by taking advantage of several methods at once [40].

From 2010, the published works introduced ML and DL to enhance the decision making to perfectly prevent 3D packing arrangements and faced related constraints in real time. Thus, from 2020 to now, the researchers are proposing new solutions to solve bin packing problem by combining previous methods and improving dynamic environment problem. In the rest of this paper, we will analyze the publications rhythm, and we will discuss the results.

To ensure the structure and quality of this review we followed the literature review process proposed by Tranfield & al. [41], which the main steps are presented in Figure 1.

3 Review: Methodology, Results

3.1 Methodology

Generally, the systematic literature review was designed in four main stages, in the 1st stage "Define" the scope of the studies is defined (as mentioned in the abstract and introduction) and a review protocol is developed in table 1. In the 2nd stage "Collect and evaluate"; the documents for research are identified and selected accordingly to criteria evaluation presented in Table 1 respecting inclusion and exclusion criteria. Only peer-review papers are considered for a reliable study. The analysis of the selected relevant papers is done on the 3rd stage "Analyze and Synthesize" with a data extraction and synthesis. The main findings are reported in the 4th stage "Reporting" and open perspectives are suggested.

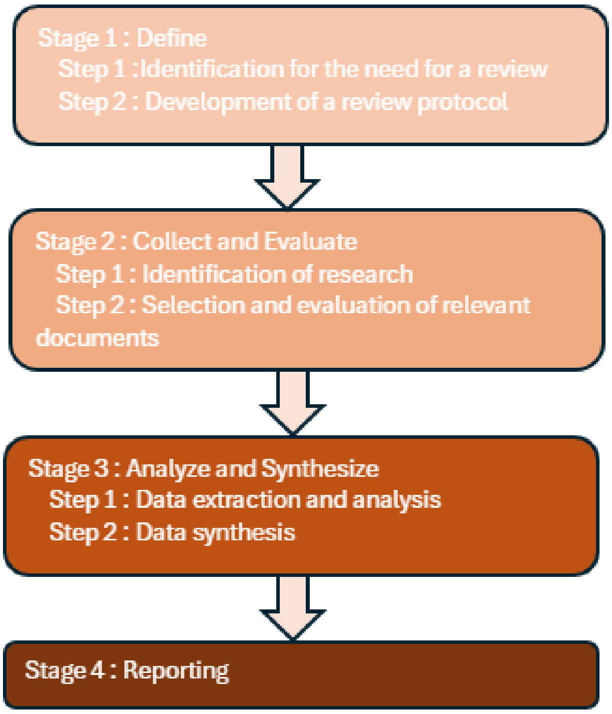


FIG. 1 : Literature review process
Source: Adapted from Tranfield et al [41]

The figure 2 presents an application of PRISMA method, designed by Page & al. [44], to our review. The number of papers retained in each step is displayed in the scheme.

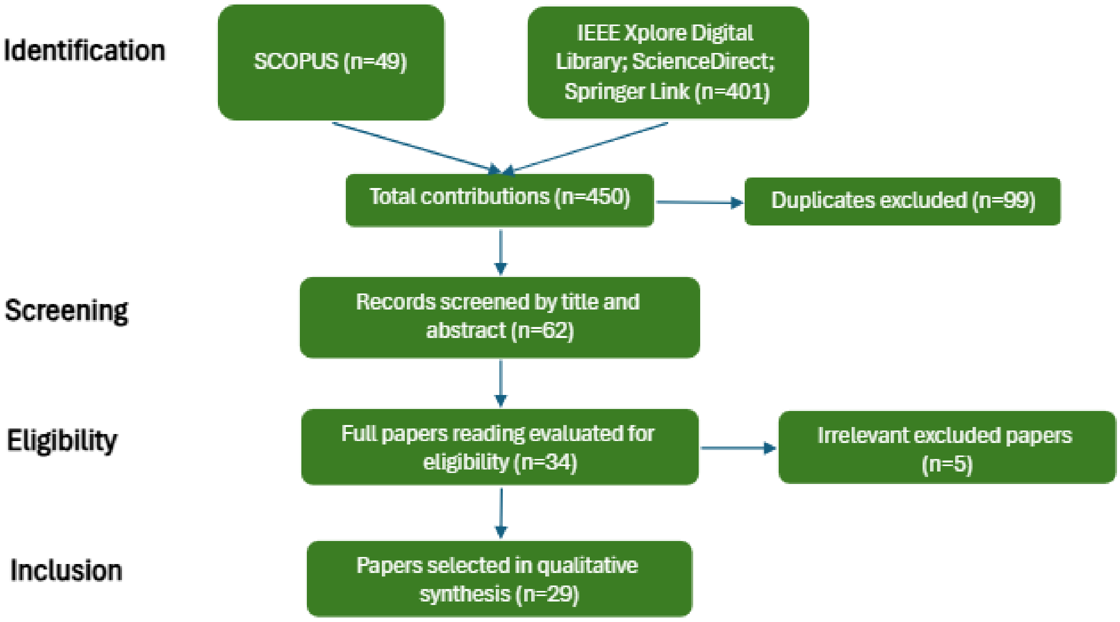


FIG. 2 : PRISMA flowchart of our review
Source: Based on Page & al. [44]

TAB. 1 : Review protocol

Research Protocol	Details description
Research various databases	Scopus is one of the largest multidisciplinary abstract and citation academic databases of peer-reviewed journals, conferences and books papers; IEEE Xplore Digital Library (IEEE); ScienceDirect (Elsevier) and Springer Link (Springer)
Publication type	Peer-reviewed papers are considered
Language	Only papers published in English were considered
Date range	The review is conducted in period between 2020 and 2024
Search fields	Titles, abstracts and keywords
Search terms	(TITLE-ABS-KEY(artificialANDintelligence)ANDTITLE-ABS-KEY(binANDpackingANDproblem)ORTITLE-ABS-KEY (machineANDlearning)ANDTITLE-ABS-KEY(binAND packingANDproblem)
Inclusion criteria	Only papers that proposed artificial intelligence (AI) and machine learning (ML) applications or a relevant discussion in the Bin packing problem were included
Exclusion criteria	Papers presenting literature reviews or papers not proposing new methods and techniques were excluded
Data extraction and analysis and synthesis	The full papers reading allow us to highlight and discuss the pertinent data and establish a global synthesis for the findings

Source: Based on Alexander & al. [42] and Pereira & al. [43]

3.2 Results

3.2.1 Publication frequency

According to the publication over a period under the heading “Artificial intelligence for bin packing problem” and “Machine learning for bin packing problem”, the annual number of papers between 2020 and 2024 is showing in the figure 3. The quantity of publications decreased after 2020, but it is notably growth after one year with a significant improvement with 10 papers in 2022 followed by 8 papers in 2023 when the papers are clearly started in decreasing.

3.1.2 The most active *databases*

Regarding the variety of papers addressing to this topic, it was noticed that almost ¼ of publications are duplications, which demonstrate that this subject is highlighted across multiple researchers and their works were published across different database.

Figure 4 illustrates the papers repartition per database, considering the duplicated ones also. Notably, Scopus has the important percentage, almost 50%, successes by IEEE Xplore Digital Library with more than ¼ of publications.

3.1.3 The most used *keywords*

VOS Viewer software tool is used in order to analyze authors keywords related to our topic “Artificial intelligence and Machine learning for bin packing problem”. A keywords map is showed in Figure 5 that illustrate clearly that the top keywords used are “Reinforcement learning”, “Machine learning”, “Heuristic algorithm” and “Optimization” followed by “Dimensional display”, “Bin packing”, “Artificial intelligence” and “Deep learning”.

4 Review: Discussion

Instead of heuristics and metaheuristics methods’ efficiency of finding optimal solutions for bin packing problem, AI is developing new technologies based on human decision and behavior. These tools, especially, ML and DL are conducting good improvement, notably:

4.1 Better optimization: AI tools such as reinforcement learning (RL) and ML play a crucial role in optimizing solutions comparing with classical heuristics [45], [46].

4.2 Computational time and cost reduction: The application of AI techniques has a notable minimization of computational time [45] and costs [47], [48].

4.3 Decision making improvement: One of the important factors of AI effectiveness is the ability of making decision based on human behavior and thinking training [48], [49].

4.4 Real time and Dynamic environment treatment: AI technologies are used to handle the real time sequential 3D bin packing problem through various methods something like a virtual dynamic environment creation [50].

4.5 Space and volume minimization: AI methods have an important contribution in space and volume optimization such as using deep learning [50], and inverse reinforcement learning [51].

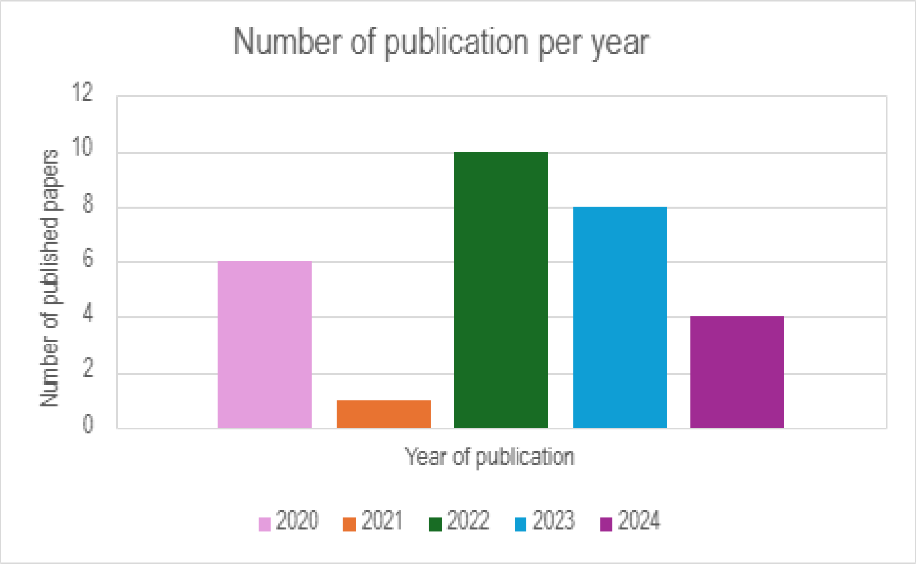


FIG. 3 : Frequency of publication per year

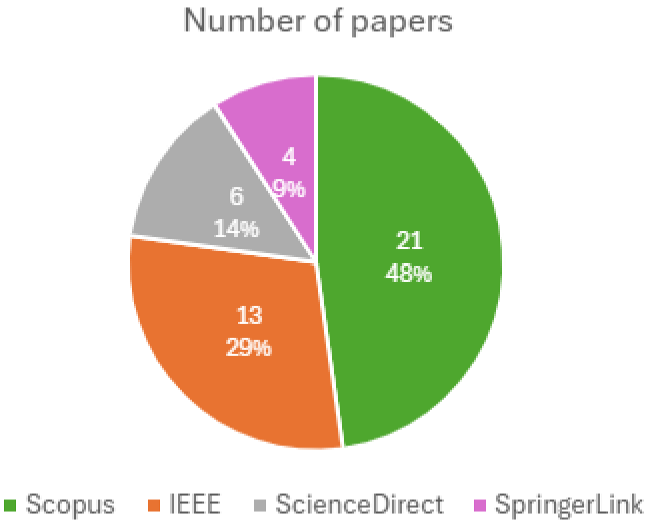


FIG. 4 : Repartition of papers per database

5 Conclusion and perspectives

Artificial intelligence and its tools, namely ML and DL contribute clearly to the development of several fields like industry, commerce, logistics, medicine, etc. This paper presents a comprehensive survey of the importance of AI in bin packing problem optimization, the results showed that in the period between 2020 and 2024, in the literature many works were published that confirmed the primordial positive impact of the related technologies on the mentioned topic in terms of quality, time and cost optimization.

The subject is clearly treated by many researchers and published in several databases; however, new applications of AI tools is highly recommended to perfectly improve new aspects of Bin Packing problem.

In the future works, the authors can propose a new hybrid algorithm combining between AI and GA and make it generic to solve a large range of real-world problem such as Bin Packing problem, Knapsack problem, Travelling Salesman problem, etc.

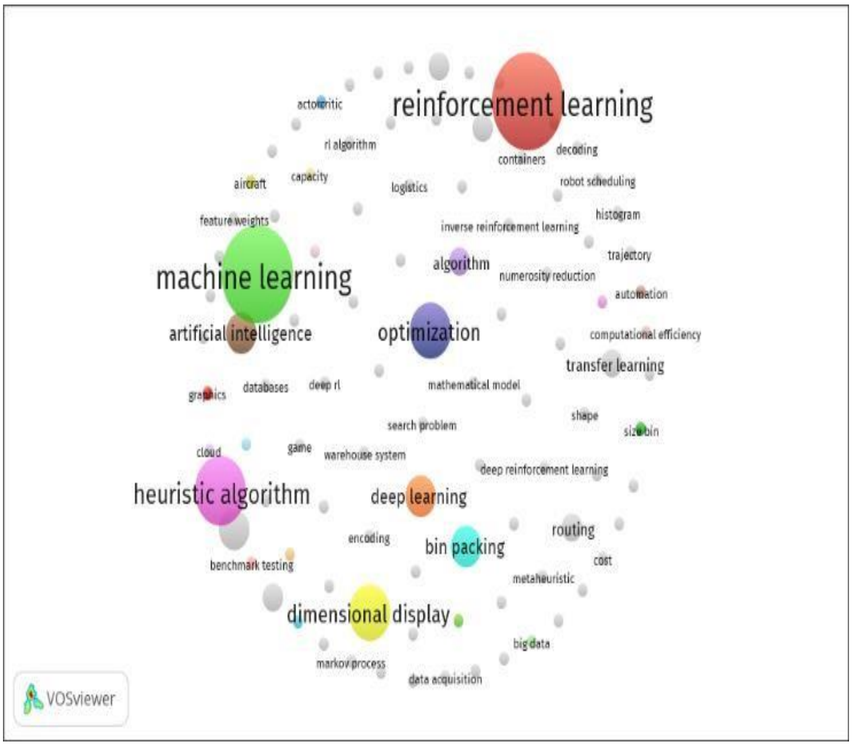


Figure 5 : Top keywords

Références

[1] Kim R.Y. (2020). The impact of COVID-19 on consumers: preparing for digital sales. IEEE Eng. Manag. Rev. doi: 10.1109/EMR.2020.2990115

[2] Martello, S., Pisinger, D., & Vigo, D. (2000). The Three-Dimensional Bin Packing Problem. Operations Research, 48(2), 256–267. doi:10.1287/opre.48.2.256.1238

[3] Pisinger D and Sigurd M (2007). Using decomposition techniques and constraint programming for solving the two-dimensional binpacking problem. INFORMS Journal on Computing 19(1):36–51

[4] Puchinger J and Raidl GR (2007). Models and algorithms for threestage two-dimensional bin packing. European Journal of Operational Research 183(3):1304–1327

[5] Sgall J. (2014). Online bin packing: Old algorithms and new results. In Conference on Computability in Europe. Cham: Springer; p. 362–72.

[6] Martello, S. & Vigo, D., (1998). Exact solution of the two-dimensional finite bin packing problem. Management science, Volume 44, pp. 388-399,

[7] Fekete SP, Schepers J. (1997). A new exact algorithm for general orthogonal d-dimensional knapsack problems. In: Algorithms ESA ’97, Springer lecture notes in computer science, vol. 1284, p. 144–56

[8] Fekete SP, Schepers J. (1997). On more-dimensional packing III: exact algorithms. Discrete Applied Mathematics, submitted for publication

[9] Berkey, J. O. & Wang, P. Y., (1987). Two-dimensional finite bin-packing algorithms. Journal of the operational research society, Volume 38, pp. 423-429

[10] Bhatia AK, Hazra M, Basu SK. (2009). Better-fit heuristic for 1D BPP. In 2009 IEEE International Advance Computing Conference; p. 193–6

[11] Chung, F. R., Garey, M. R. & Johnson, D. S., (1982). On packing two-dimensional bins. SIAM Journal on Algebraic Discrete Methods, Volume 3, pp. 66-76

[12] Lodi A, Martello S and Vigo D. (2002). Heuristic algorithms for the three-dimensional bin packing problem. European Journal of Operational Research 141(2):410–420

[13] Monaci, M., & Toth, P. (2006). A set-covering-based heuristic approach for bin-packing problems. INFORMS Journal on Computing, 18(1), 71–85, Z

[14] Z. Dai, and J. Cha, (1994). “A Hybrid Approach of Heuristic and Neural Network for Packing Problems,” Advances in Design Automation 1994: Proceedings of the 20th ASME Design Automation Conference, vol. 2, pp 117-123

[15] Puchinger, J. & Raidl, G. R., (2004). An evolutionary algorithm for column generation in integer programming: an effective approach for 2D bin packing. In International Conference on Parallel Problem Solving from Nature. In International Conference on Parallel Problem Solving from Nature, September, pp. 642-651

[16] Schultz KM, Passino KM, Seeley TD. (2008). The mechanism of flight guidance in honeybee swarms: subtle guides or streaker bees?. J Exp Biol. 211(20):3287–95

[17] Soke, A. & Bingul, Z., (2006). Hybrid genetic algorithm and simulated annealing for two-dimensional non-guillotine rectangular packing problems. Engineering Applications of Artificial Intelligence, Volume 19, pp. 557-567

[18] Mirjalili S, Lewis A. (2016). Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:51–67

[19] Perboli G, Crainic TG and Tadei R (2011). An efficient metaheuristic for multi-dimensional multi-container packing. In: 2011 IEEE Conference on Automation Science and Engineering (CASE). Presented at the 2011 IEEE Conference on Automation Science and Engineering (CASE), 563–568

[20] A.Moura and J. F. Oliveira. (2005). "A GRASP approach to the Container-Loading Problem," Intelligent Systems, IEEE, vol. 20, pp. 50-57

[21] Maple, C., Szpruch, L., Epiphaniou, G., Staykova, K., Singh, S., Penwarden, W., Wen, Y., Wang, Z., Hariharan, J. and Avramovic, P., (2023). The ai revolution: opportunities and challenges for the finance sector, https://doi.org/10.48550/arXiv.2308.16538

[22] Saxena, Ashish K. and Ness, Stephanie and Khinvasara, Tushar (2024) The Influence of AI: The Revolutionary Effects of Artificial Intelligence in Healthcare Sector. Journal of Engineering Research and Reports, 26 (3). pp. 49-62. ISSN 2582-2926

[23] L. Krstić, V. Aleksić, and M. Krstić. (2022). “Artificial Intelligence in Education: A Review,” in Proceedings TIE 2022, University of Kragujevac, Faculty of Technical Sciences Čačak, pp. 223–228. doi: 10.46793/TIE22.223K

[24] Ahmed, I., Jeon, G., & Piccialli, F. (2022). From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where. IEEE Transactions on Industrial Informatics, 18(8), 5031-5042. https://doi.org/10.1109/TII.2022.3146552

[25] Richey Jr, R. G., Chowdhury, S., Davis-Sramek, B., Giannakis, M., & Dwivedi, Y. K. (2023). Artificial intelligence in logistics and supply chain management: A primer and roadmap for research. Journal of Business Logistics, 44(4), 532- 549

[26] Garey MR and Johnson DS. Garey MR and Johnson DS. (1979). Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman.

[27] Sridhar, R., Chandrasekaran, M., Sriramya, C., & Page, T. (2017). Optimization of heterogeneous Bin packing using adaptive genetic algorithm. IOP Conference Series: Materials Science and Engineering, 183, 012026. doi:10.1088/1757-899x/183/1/012026

[28] Munien, C., & Ezugwu, A. E. (2021). Metaheuristic algorithms for one-dimensional bin-packing problems: A survey of recent advances and applications. Journal of Intelligent Systems, 30(1), 636–663. doi:10.1515/jisys-2020-0117

[29] Gilmore, P. C., & Gomory, R. E. (1965). Multistage cutting stock problems of two and more dimensions. *Operations research*, *13*(1), 94-120.

[30] Egeblad, J., & Pisinger, D. (2009). Heuristic approaches for the two- and three-dimensional knapsack packing problem. Computers & Operations Research, 36(4), 1026–1049. doi:10.1016/j.cor.2007.12.004

[31] Chen, C. S., Lee, S. M., & Shen, Q. S. (1995). *An analytical model for the container loading problem. European Journal of Operational Research, 80(1), 68–76.* doi:10.1016/0377-2217(94)00002-t

[32] McCarthy, J.; Minsky, M.L.; Rochester, N.; Shannon, C.E. (1955). A proposal for the dartmouth summer research project on artificial intelligence. AI Mag. 2006, 27, 12

[33] Farzaneh, H., Malehmirchegini, L., Bejan, A., Afolabi, T., Mulumba, A., & Daka, P. P. (2021). Artificial Intelligence Evolution in Smart Buildings for Energy Efficiency. Applied Sciences, 11(2), 763. doi:10.3390/app11020763

[34] Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). Optimization by simulated annealing. science, Volume 220, pp. 671- 680

[35] Drexl, A. (1988). A simulated annealing approach to the multiconstraint zero-one knapsack problem. *Computing*, *40*(1), 1-8.

[36] Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. ann arbor: University of michigan press. *Ann Arbor: The University of Michigan Press*.

[37] Khuri, S., Bäck, T. & Heitkötter, J., (1994). The Zero/One Multiple Knapsack Problem and Genetic Algorithms. Proceedings of the ACM Symposium on Applied Computing, pp. 188-193

[38] Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), 26(1), 29–41. doi:10.1109/3477.484436

[39] Kennedy, J., & Eberhart, R. (n.d.). (1995). Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks. doi:10.1109/icnn.1995.488968

[40] Karabulut, K., & İnceoğlu, M. M. (2004). A Hybrid Genetic Algorithm for Packing in 3D with Deepest Bottom Left with Fill Method. Lecture Notes in Computer Science, 441–450. doi:10.1007/978-3-540-30198-1_45

[41] Tranfield, D., Denyer, D. and Smart, P. (2003), “Towards a methodology for developing evidence-informed management knowledge by means of systematic review”, British Journal of Management, Vol. 14 No. 3, pp. 207-222

[42] Alexander, A., Walker, H. and Naim, M. (2014), “Decision theory in sustainable supply chain management: a literature review”, Supply Chain Management: An International Journal, Vol. 19 Nos 5/6, pp. 504-522

[43] Pereira, C.R., Christopher, M. and Lago Da Silva, A. (2014), “Achieving supply chain resilience: the role of procurement”, Supply Chain Management: An International Journal, Vol. 19 Nos 5/6, pp. 626-642

[44] M. J. Page et al. (2021). “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” Int. J. Surg., vol. 88, p. 105906, 2021

[45] Fang, Jie & Rao, Yunqing & Zhao, Xusheng & Du, Bing. (2023). A Hybrid Reinforcement Learning Algorithm for 2D Irregular Packing Problems. Mathematics. 11. 327. 10.3390/math11020327.

[46] Guerriero, Francesca & Saccomanno, Francesco Paolo. (2023). A Machine Learning Approach for the Bin Packing Problem. 436-441. 10.1109/IDAACS58523.2023.10348703.

[47] Singh, Emilio & Pillay, Nelishia. (2022). A Study of Transfer Learning in an Ant-Based Generation Construction Hyper-Heuristic. 1-7. 10.1109/CEC55065.2022.9870415.

[48] Bruni, Maria & Fadda, Edoardo & Fedorov, Stanislav & Perboli, Guido. (2023). A machine learning optimization approach for last-mile delivery and third-party logistics. Computers & Operations Research. 157. 106262. 10.1016/j.cor.2023.106262.

[49] Danino, Tom & Ben-Shimol, Yehuda & Greenberg, Shlomo. (2023). Container Allocation in Cloud Environment Using Multi-Agent Deep Reinforcement Learning. Electronics. 12. 2614. 10.3390/electronics12122614.

[50] Murdivien, Shokhikha & Um, Jumyung. (2023). BoxStacker: Deep Reinforcement Learning for 3D Bin Packing Problem in Virtual Environment of Logistics Systems. Sensors. 23. 6928. 10.3390/s23156928.

[51] H. Foot and B. Mättig, (2024). "Towards Adapting Human Behaviour in Packing Logistics Through Imitation Learning," 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Osaka, Japan , pp. 331-336, doi: 10.1109/ICAIIC60209.2024.10463496.

Blockchain Technology in Supply Chain Finance: An In-Depth Review and Future Research Directions

HALIMA SEMAA
University moulay ISMAIL
ERRACHIDIA , MOROCCO
halimasemaa9@gmail.com

YOUSSEF MALHOUNI
UNIVERSITY HASSAN PREMIER
SETTAT,MOROCCO
youssefmalhouni@gmail.com

LAILA BOUZARRA
university moulay ismail
ERRACHIDIA ,MOROCCO
bouzarra03@gmail.com

HANAE KERROUCH
University moulay ISMAIL
ERRACHIDIA, MOROCCO
Hanae.kerrouch@gmail.com

MOHAMED AIT HOU
UNIVERSITY MOULAY ISMAIL
ERRACHIDIA ,MOROCCO
m.aithou@fpe.umi.ac.ma

Abstract

In recent years, blockchain technology has garnered increasing interest due to its significant potential for improving information management. Particularly, research has increasingly focused on its application in the supply chain finance (SCF) domain. This paper provides an in-depth literature review, exploring various ways blockchain technology is used in SCF. We identify and specify the key research themes in this area, highlighting its implications for optimizing financial flows, reducing information asymmetry, and securing transactions within supply chain networks.

1. Introduction

This concept (SCF), which has emerged recently [1] aims to optimize financial flows between organizations [2]. It is based on a set of solutions developed by financial institutions in collaboration with major buyers, aiming to provide partner companies with effective means to optimize their cash flows, improve visibility on commercial flows, and access alternative financing under more advantageous conditions than those obtained directly [3,4]. The primary objective is to align financial flows with physical and information flows within the supply chain, thereby enhancing cash flow management.toutefois , plusieurs entreprises sont confronté à differents defis dans l'application de la SCF .En effet,

With the credit approval of large enterprises, financial institutions such as banks are willing to provide financing services such as accounts receivable financing, factoring, inventory financing, etc., to primary suppliers/distributors. However, many secondary and beyond suppliers/distributors struggle to obtain financing due to their lack of direct connection with large enterprises. In practice, risk management in SCF primarily focuses on financing applicants and their respective transactions. Nevertheless, verifying the authenticity of involved parties, transactional data, and underlying assets is often challenging. Moreover, the diversity of participants and the complexity of business processes make risk management more challenging than in traditional financial transactions.pour resoudre ces probleme, La technologie blockchain provides a solution to address information barriers among SCF enterprises. As a shared and tamper-proof database, blockchain not only ensures information security and prevents information asymmetry and fraud within the supply chain, but also enhances the trust relationship between businesses and financial institutions. In this paper, we will conduct a literature review on the application of blockchain technology in supply chain finance, focusing on key research themes.

The first section will introduce the fundamental concepts of SCF and blockchain. The second section will discuss key themes explaining the application of blockchain in SCF.

2. General concepts

1.1 Supply chain finance

Supply chain finance (SCF) is a comprehensive set of financial solutions designed to optimize cash flow management and enhance the liquidity of businesses involved in a supply chain. These solutions facilitate the smooth operation of financial transactions and help maintain a healthy cash flow throughout the supply chain. SCF relies on the collaboration between various entities such as suppliers, manufacturers, distributors, retailers, and financial institutions. This collaboration ensures that financial operations related to commercial transactions are managed efficiently and effectively. One of the primary goals of SCF is to optimize the need for working capital [2 ,3,4,5]. Working capital is crucial for day-to-day business operations as it covers expenses such as salaries, raw materials, and other operational costs. SCF helps achieve this by improving the management of receivables (money owed to the business), payables (money the business owes), and inventory (goods the business holds for sale). By employing various SCF solutions such as invoice financing, factoring, reverse factoring, and dynamic discounting, businesses can accelerate the conversion of receivables into cash, extend payment terms with suppliers without straining their relationships, and optimize inventory levels to free up cash. For instance, invoice financing allows businesses to obtain immediate funds by selling their outstanding invoices to financial institutions at a discount. This provides quick access to cash that can be reinvested into the business. Similarly, reverse factoring enables suppliers to receive early payments for their invoices, improving their cash flow while allowing buyers to maintain longer payment terms. Dynamic discounting allows buyers to offer early payments to suppliers in exchange for discounts, benefiting both parties by improving liquidity and reducing costs [6,21].

1.2 Blockchain technology

Satoshi Nakamoto (2008) introduced the digital currency "Bitcoin" in a paper titled "Bitcoin: A Peer-to-Peer Electronic Cash System" [8]. This paper marks the first mention of blockchain technology (BCT). Blockchain serves as a sequential database that links blocks of data using timestamp technology. It functions as a tamper-resistant infrastructure for contract management and a distributed computing paradigm, employing a chain structure to verify and store data. Blockchain utilizes a distributed consensus mechanism for appending and updating data and supports scripting code (smart contracts) to program and manipulate ledger records [9]. Blockchains are generally categorized into two types: permissioned and unpermissioned. The unpermissioned blockchain, like the public blockchain (e.g., Bitcoin), allows each node equal access and privileges on the ledger. Permissioned blockchains, such as consortium and private blockchains, restrict ledger operations to authorized nodes only. In its early phases, Blockchain enabled direct Bitcoin payments among parties distrustful of each other, bypassing the need for intermediaries [9]. Its decentralization and ability to facilitate cross-border payments have significantly impacted traditional finance. Vitalik Buterin (2014) introduced Ethereum, which integrates smart contracts with blockchain technology (BCT), enabling users to create and execute smart contracts and develop diverse decentralized applications [10]. Swan (2015) argued that BCT has the potential to fundamentally transform traditional credit systems and become pivotal in the fourth industrial revolution [11]. BCT-based systems have the capacity to streamline the financial industry by eliminating inefficiencies, high costs associated with manual operations, complex processes, and inconsistent standards, thereby bringing disruptive changes. For instance, blockchain's traceability feature finds extensive applications in supply chain management, logistics, and the Internet of Things (IoT) [12–13].

1.3 Blockchain technology in a supply chain

Global changes, digital evolution, and financial developments significantly impact the profitability of the entire supply chain [14]. Blockchain technology has the potential to be considered one of the most effective transformative technologies. Blockchain technology (BT) can be applied in records management, public asset management, financial institution ledger execution, and financial asset settlement [14].

By implementing a blockchain-enabled supply chain, companies can expand a real-time distributed ledger of purchases and interactions for all partners in their supply chain network. BT helps companies make reliable demand forecasts, effectively manage capital, and reduce product transportation costs. At the same time, it provides clear visibility to financing companies on their investment in this financing through its ability to generate activity logs. In summary, this can help stakeholders reduce risks at a lower cost compared to the traditional method, which requires excess capacity and third-party backups.

3. Related review

1.4 Using Blockchain and IoT to Enable Transparency of Information in Supply Chain Financing

In an SCF (Supply Chain Finance) transaction, numerous transaction records such as orders, customer accounts, invoices, supplier accounts, and financial guarantees are generated among different parties. Typically, these records are created bilaterally and are only accessible to the directly involved parties. For instance, ERP (Enterprise Resource Planning) systems are generally not accessible to other businesses in the supply chain, complicating access to comprehensive information for indirectly involved parties like upstream and downstream SMEs. This leads to an "information silo" issue where commercial information between SMEs and principal enterprises is not effectively tracked. Moreover, SMEs often lack the necessary IT capabilities to digitize internal management or automate external collaboration with business partners, making it challenging for them to integrate into the supply chain ecosystem [15].

L.Guo et al. (2022) proposed a framework for information management based on BCT and IoT, named BC4Regu, which acts as a regulator to improve information transparency in SCF business processes. This platform facilitates effective regulation of information flow and business processes, enhancing operational management and reducing risks and costs in SCF. By leveraging a distributed ledger, BC4Regu creates a tamper-resistant digital asset management system. The integration of smart contracts effectively optimizes traditional financing processes. The authors found that the integration of blockchain technology (BCT) into supply chain financing activities can effectively encourage SMEs to improve their operational capabilities to increase their revenues. At the same time, the principal can benefit from this improvement by providing SMEs with blockchain technology and incentive subsidies [15].

Teng et al. (2024) studied the diffusion mechanism of blockchain technology among upstream and downstream companies, analyzing the conditions influencing this diffusion to address the issue of information asymmetry and enhance the efficiency of supply chain finance (SCF). This study demonstrates the diffusion mechanism of blockchain technology in SCF and analyzes the stability of this diffusion model using the Lyapunov method, drawing on diffusion theory impacting technological innovation and the contagion model describing the spread of diseases [16].

Wang et al (2023) . highlighted that distrust, privacy concerns, data misuse, or asymmetries frequently hinder data sharing among entities. They proposed using blockchain technology to overcome these challenges, enabling businesses to establish an effective, fair, secure, and trustworthy data sharing mechanism [17].

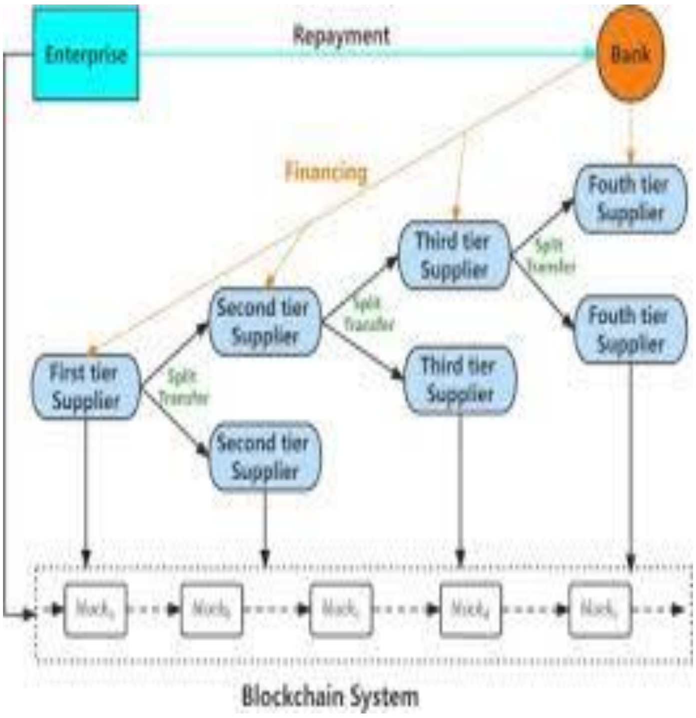


FIG.1– A typical Blockchain-based Supply Chain Finance Scenario (Hu et al 2023)

The author Hu et al 2023 has proposed an adaptive consortium blockchain sharding framework named ACSarF, based on a deep reinforcement learning algorithm, to address the problem of static sharding, which lacks adaptability for the dynamic SCF environment, or those designed for public chains, making them inapplicable to consortium blockchain-based SCF. The proposed framework can improve consortium blockchain sharding by effectively reducing transaction delays and adaptively adjusting the sharding and block production strategies to increase the transaction success rate in a dynamic SCF environment. Additionally, they have proposed using a consistent hashing algorithm within the ACSarF framework to ensure transaction load balancing in the adaptive sharding system, further enhancing the performance of blockchain sharding in dynamic SCF scenarios. [18].

1.5 Identifying the Determinants of Blockchain Application in Supply Chain Financing

In another study, the author Kabir, M. R et (2021) explored the determining factors influencing the adoption of blockchain technology in supply chain financing platforms. This study aims to analyze these determinants such as Performance Expectancy (PEXP), Effort Expectancy (EEXP), Social Influence , Trust, Facilitating Conditions (FCON) , and Behavioural Intention ((BINTU) using the Unified Theory of Acceptance and Use of Technology (UTAUT). UTAUT is a relevant approach to understand how the adoption of technologies like blockchain for supply chain financing is perceived and implemented [15]. The study validated the hypotheses and indicates that the predictive capability of FCON and EEXP suggests that financial institutions and their clients have a favorable attitude towards adopting modern technologies such as blockchain for their supply chain financing operations, provided they receive organizational support from service providers and ensure smooth and efficient operations. Similarly, PEXP positively influences users' intention to adopt blockchain, as they believe this technology can enhance their performance [19].

1.6 Use of blockchain for sharing corporate credit information in the supply chain.

To address the financing needs of businesses within the supply chain, traditional SCM management faces significant challenges. Many companies encounter difficulties in SCF due to their liquidity needs in an environment where available credit is insufficient. This situation complicates credit transfers between major suppliers and SMEs, thereby exacerbating their financial problems. This study aims to enhance data sharing and blockchain-based traceability of chain blocks through improved consensus mechanisms, aiming to strengthen the existing credit reporting services system. The author proposes designing a credit investigation system for SCF based on blockchain technology, with the potential goal of addressing the dilemma faced by supply chain enterprises - the need for funds despite a lack of available credit. This approach holds significant theoretical and practical importance [20].

Zheng et al. developed a blockchain-based model for access control and the management of shared transaction information, which optimizes the current credit system and increases its efficiency.Chod et al. demonstrated that information transparency enhances the ability of companies to obtain favorable financing conditions with reduced information transfer costs. Regarding information sharing, blockchain technology reduces the risk of default in supply chain finance (SCF). Babich and Hilary stated that the application of blockchain technology to SCF can overcome information asymmetry and decrease the costs associated with default risks caused by ethical issues [20].

4. Conclusion and Future research direction

Currently, blockchain technology applied to supply chain finance is in its early exploratory stages. There are few studies addressing this topic. These studies have demonstrated the importance of blockchain technology in improving SCF practices. According to the research cited above, blockchain helps reduce information asymmetry and prevents data falsification in the SCF credit system. It facilitates information sharing, especially for SMEs lacking a robust credit system. However, legal issues (e.g., the legal validity of proof of ownership of asset tokens and smart contracts themselves) hindering widespread adoption of blockchain technology in SCF have not been explored in these studies, which presents a research avenue. Additionally, these studies diverge from various psychosocial assumptions, such as evaluation mechanisms for protection, risks, and security for applications, connectivity demand, and presumed correlation among user groups and other implementation-related aspects of blockchain technology. Moreover, the influence of time intervals in the technology diffusion process is overlooked in some studies.

5. References

[1] Gelsomino, L. M., Mangiaracina, R., Perego, A., & Tumino, A. (2016). Supply chain finance: a literature review. International Journal of Physical Distribution & Logistics Management, 46(4).

[2] Hofmann, E. (2005). Supply chain finance: some conceptual insights. Beiträge Zu Beschaffung Und Logistik, 16(1), 203-214.

[3] Camerinelli, E. (2009). Supply chain finance. Journal of Payments Strategy & Systems, 3(2), 114-128.

[4] Lamoureux, J. F., & Evans, T. A. (2011). Supply chain finance: a new means to support the competitiveness and resilience of global value chains. Available at SSRN 2179944.

[5] Hofmann, E. (2011). Supply chain finance solutions. Springer.

[6] Pfohl, H. C., & Gomm, M. (2009). Supply chain finance: optimizing financial flows in supply chains. Logistics research, 1, 149-161.

[7] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

[8] K. Christidis, M. Devetsikiotis, Blockchains and smart contracts for the internet of things, IEEE Access 4 (2016) 2292–2303.

[9] V. Buterin, A Next-Generation Smart Contract and Decentralized Application Platform, White Paper, 2014. https://www.blockchainlab.com/pdf/Ethereum_wh ite_paper-a_next_generation_smart_contract_and_decentralized_application_platfo rm-vitalik-buterin.pdf/

[10] Swan, M. (2015). Blockchain: Blueprint for a new economy. " O'Reilly Media, Inc.".

[11] K. Yan, A. Chong, Y. Mo, Generative Adversarial Network for Fault Detection Diagnosis of Chillers, Building and Environment, 2020, p. 172.

[12] S.M. Mohammad, Z.H. Joshua, S. Salman, Z. Tamer, E.K. Sadatdiynov, A survey of data partitioning and sampling methods to support big data analysis, Big Data Mining and Analytics 3 (2) (2020) 85–101.

[13] IEEE Softw 34 (6) (2017) 21–27. [11] I. Eyal, Blockchain technology:

transforming libertarian cryptocurrency dreams to finance and banking realities, Computer 50 (9) (2017) 38–49

[14] Guo, L., Chen, J., Li, S., Li, Y., & Lu, J. (2022). A blockchain and IoT-based lightweight framework for enabling information transparency in supply chain finance. Digital Communications and Networks, 8(4), 576-587.

[15] Guo, L., Chen, J., Li, S., Li, Y., & Lu, J. (2022). A blockchain and IoT-based lightweight framework for enabling information transparency in supply chain finance. Digital Communications and Networks, 8(4), 576-587.

[16] Teng, Y., Ma, S., Qian, Q., & Wang, G. (2024). SEIR-diffusion modeling and stability analysis of supply chain finance based on blockchain technology. Heliyon, 10(3).

[17] Wang, L., Zhang, J., Li, J., Yu, H., & Li, J. (2023). An ISM-DEMATEL analysis of blockchain adoption decision in the circular supply chain finance context. Management Decision, (ahead-of-print).

[18] Hu, S., Lin, J., Du, X., Huang, W., Lu, Z., Duan, Q., & Wu, J. (2023). ACSarF: a DRL-based adaptive consortium blockchain sharding framework for supply chain finance. Digital Communications and Networks.

[19] Kabir, M. R., Islam, M. A., Marniati, & Herawati. (2021). Application of blockchain for supply chain financing: explaining the drivers using SEM. Journal of Open Innovation: Technology, Market, and Complexity, 7(3), 167.

[20] Zheng, K., Zheng, L. J., Gauthier, J., Zhou, L., Xu, Y., Behl, A., & Zhang, J. Z. (2022). Blockchain technology for enterprise credit information sharing in supply chain finance. Journal of Innovation & Knowledge, 7(4), 100256.

[21] Caniato, F., Henke, M., & Zsidisin, G. A. (2019). Supply chain finance: historical foundations, current research, future developments. Journal of Purchasing and Supply Management, 25(2), 99-104.

Comparative Analysis of Genetic Algorithm and Reinforcement Learning for solving the Vehicle Routing Problem

Achraf Berrajaa(1), Naoufal ROUKY(2), Abdelhamid Benaini(3) and Jaouad Boukachour(3)

(1) Department of computer Science, Faculty of Sciences, University Mohamed First, Oujda, Morocco
(2) Artificial Intelligence Research and Applications Laboratory , Faculty of Science and Technology, Hassan First University, Settat 26000, Morocco
(3) Normandie University, LMAH, 76063 le Havre, France.

Abstract. This research proposes a comprehensive comparative analysis of two leading optimization methodologies, namely, the Genetic Algorithm (GA) and Reinforcement Learning (RL), applied to solve the Vehicle Routing Problem (VRP). The VRP, a complex combinatorial optimization challenge, holds significant practical applications, particularly in logistics and transportation management.

The study examines and contrasts the effectiveness of GA and RL in optimizing routes for a fleet of vehicles tasked with serving a set of geographically dispersed locations. The comparison of performance between these two algorithms spans criteria such as solution quality, computational efficiency, and scalability. These results provide meaningful insights into the respective strengths and limitations of GA and RL in addressing the VRP. This comparative analysis aims to offer valuable guidance for practitioners and researchers facing similar combinatorial optimization challenges across various domains.

Keywords: Artificial intelligence; Reinforcement Learning; Genetic Algorithm; Vehicle Routing Problem.

1 Introduction

The Vehicle Routing Problem (VRP) stands out as a critical NP-hard combinatorial optimization challenge, prompting the development of numerous heuristic methods. In this paper, we delve into the fundamental VRP, a task involving several vehicles with a fixed capacity C tasked with delivering known quantities qi of goods to n cities from a single depot. Considering the distances d_{ij} between cities, the objective is to devise optimal tours for the vehicles, ensuring that each city is visited exactly once, and the total quantity carried on any tour does not surpass the vehicle capacity C [2]. This classic problem has given rise to various VRP variants, such as those incorporating time windows, fleet size and mix vehicles, split delivery, and more recent endeavors into electric and green VRP. The extensive literature on VRP methodologies commenced with exact approaches like linear programming, dynamic programming, and branch-and- bound algorithms due to its NP-hard nature. However, the richness of heuristic and metaheuristic solutions has become prominent, including methods like the Tabu approach, Genetic Algorithms, and Ant Colony Algorithms. With the evolution of artificial intelligence, it is now opportune to conduct a comparison between two models regarded as artificial intelligence models. This chapter introduces a comprehensive comparative analysis of two prominent optimization methodologies: Genetic Algorithms (GAs) and Reinforcement Learning (RL). The emphasis is on their application to solve the Vehicle Routing Problem (VRP), a task that entails determining optimal routes for a fleet of vehicles to service a set of customers with known demands while minimizing overall costs.

The rest of this chapter is structured as follows: Section 2 provides a review of solutions and technologies presented in the literature for solving the VRP. Section 3 outlines the considered problem, detailing its parameters and constraints. In Section 4, we delve into the adopted research methodology, presenting the proposed architecture for our Genetic Algorithm (GA) and outlining the steps for our Reinforcement Learning (RL). Section 5 is dedicated to presenting experimental results, accompanied by a comparative analysis between the two algorithmic approaches. Finally, Section 6 summarizes all findings, providing a conclusion to the work undertaken and outlining future perspectives for this research.

2 Literature review

Early research on the Vehicle Routing Problem (VRP) mainly focused on traditional optimization techniques and exact algorithms. These methods included linear programming [7], dynamic programming [8], and branch-and-bound algorithms [9]. Although effective for small-scale instances, these exact methodologies face challenges when dealing with larger problem sizes due to the NP-hard nature of the VRP. As a result, attention shifted towards heuristic and metaheuristic approaches that could efficiently handle the complexity of real-world routing scenarios. Given the computational intractability of the VRP, numerous heuristic and metaheuristic methods have been developed to find near optimal solutions within reasonable time frames [4]. Notable among these are Tabu methods [1], Genetic Algorithms (GAs) [3], and Ant Colony Optimization (ACO) [10, 11]. These methods leverage intelligent search strategies to explore the solution space effectively and quickly converge to satisfactory solutions. GAs, inspired by the process of natural selection, use genetic operators such as crossover and mutation to evolve a population of candidate solutions. Genetic Algorithms have demonstrated remarkable success in solving complex optimization problems by mimicking the principles of natural selection. In the context of the VRP, GAs evolve a population of potential solutions over generations, with each solution representing a unique routing plan. This iterative process involves the crossover and mutation of candidate solutions, gradually refining the population to converge towards an optimal or near optimal solution. Among the most interesting contributions, we mention [3] where the authors proposed an innovative contribution in parallel with a Genetic Algorithm (GA) architecture, leveraging the capabilities of Graphics Processing Units (GPUs) to accelerate convergence. The adaptability of GAs makes them particularly well-suited for handling the intricacies of the VRP, where route optimization involves balancing multiple constraints, including vehicle capacity and time window considerations.

Furthermore, with the advent of metaheuristics and heuristics, the progress of artificial intelligence has opened new perspectives for solving the Vehicle Routing Problem (VRP) using intelligent methods, notably reinforcement learning.

Reinforcement Learning (RL) has proven to be a promising approach to enhance the efficiency of VRP resolution iterations by enabling algorithms to learn and adapt to specific problem models over time. Reinforcement Learning (RL) represents a paradigm shift in optimization strategies. RL agents learn optimal behavior through interaction with an environment, receiving feedback in the form of rewards or penalties based on their actions. When applied to the VRP, RL models dynamically learn to adapt vehicle routes in response to changing conditions, making them particularly relevant for dynamic or real-time VRP scenarios. The flexibility of RL to adapt to changing environments and learn from experience positions, it as a promising approach to address the inherent complexities of VRP. A notable study in this context is that of Lin et al. [6], where the authors introduced a reinforcement learning-based approach to solve the VRP with time windows. In addition, in [5], the authors proposed a contribution based on RL to solve the dynamic vehicle routing problem with stochastic customers.

In this chapter, we intend to provide a comparison between our two approaches, the first based on a Genetic Algorithm (GA) and the second on Reinforcement Learning (RL). This comparative analysis aims to provide insights into the strengths and weaknesses of Genetic Algorithms and Reinforcement Learning when tackling VRPs. By evaluating their performance on standard VRP benchmarks, we aim to shed light on the specific contexts in which each methodology excels. Understanding the trade-offs and considerations associated with GAs and RL in the VRP domain is crucial for practitioners seeking to implement effective and tailored optimization solutions for their specific routing challenges.

3 Problem Statement

The Vehicle Routing Problem (VRP) is a classic combinatorial optimization problem that focuses on determining the most efficient routes for a fleet of vehicles to serve a set of customers, subject to specific constraints. Formally, the VRP involves a set of N nodes, representing customers, and a central depot. Each customer has a demand q_i, which specifies the quantity of goods to be delivered. The fleet comprises V vehicles, each with a uniform capacity C. The objective of the VRP is to minimize the total distance traveled or the total cost of delivery, while ensuring the following constraints are satisfied.

1. The total demand on any given route does not exceed the vehicle's capacity C.
2. All customers are served exactly once by one of the vehicles.
3. Each vehicle starts and ends its route at the depot.

Figure 1 gives an example of VRP with n = 10 customers to serve. Node numbers are written inside the routes and the demands to the outside.

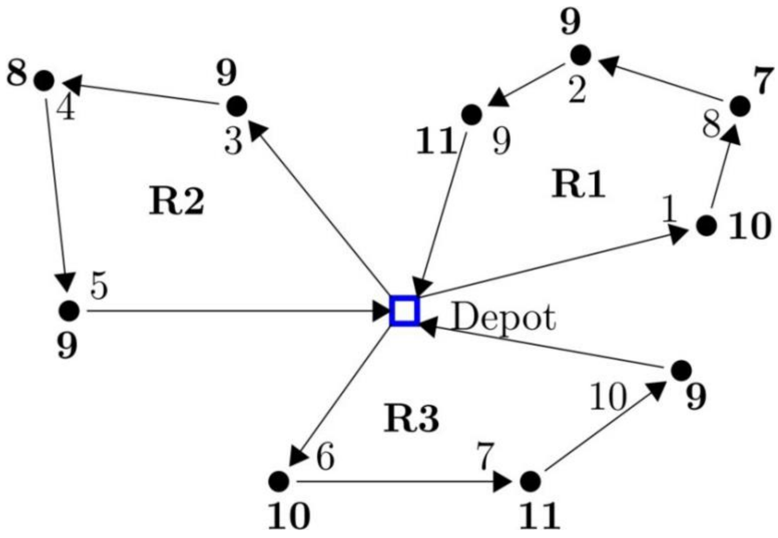


FIG .1 – Example of a feasible VRP solution with a maximum vehicle capacity *C=40*

4 The proposed contribution

As mentioned earlier, we will introduce two algorithms to address the Vehicle Routing Problem (VRP). The first algorithm is the Genetic Algorithm (GA), while the second one relies on Reinforcement Learning (RL).

4.1 Genetic Algorithm

The Genetic Algorithm is a type of evolutionary method that simulates the natural process of biological evolution. In GA, a new population is generated from the previous population using genetic operators iteratively, typically with a stop after a specified number of iterations. A crucial aspect of GA is the selection of solution representation, also known as chromosome encoding. In this context, we adopt the Set-Based Representation introduced in [3] to encode a solution for the VRP as follows. Each route, Ri, in the solution is represented by the ordered set of nodes (cluster) Ci that make up the route. The entire solution is encoded by the clusters C1, · · · ,Cm. This encoding explicitly outlines the solution's routes, eliminating the need for a decoding step. For example, the encoding of the solution in Figure 1 is represented as C1 = (0, 1, 8, 2, 9, 0), C2 = (0, 3, 4, 5, 0) and C3 = (0, 6, 7, 10, 0).

We define straightforward genetic operators, which are applied iteratively to the encoding of an initial solution with the aim of enhancing its quality.

a. **Crossover Operator**

The crossover operation stands as the primary genetic operator within a Genetic Algorithm, and it is inherently complex to articulate. The process of crossing two chromosomes, *p* and *p′*, is delineated through the following four steps:

Step 1: Randomly select a cluster C from p, and another cluster C' from p'.

Step 2: Execute a random cut at two points within the chosen clusters C and C'. Denote X (respectively X') as the set of nodes situated between the two cut points of C (respectively C'). Essentially, X (respectively X') constitutes the cut in C (respectively C'). The length (number of nodes) of cuts X and X' must be equivalent. Furthermore, the initial nodes of C and C' should not be included in the cuts. To maintain simplicity, we restrict the cut length to 1, 2, or 3 nodes, allowing for the interchange of a single node, a two-node arc, or a three-node path between C and C'.

Step 3: Substituting X with X' in C and X' with X in C' is performed, provided that this substitution does not violate the capacity constraint, resulting in two offspring solutions, *ch* and *ch′*. If a violation occurs, the process returns to Step1.

Step 4: Nodes in $Y = X \cap X' \setminus X'$ appear twice in *ch'*, while nodes in $Y' = X' \cap X \setminus X'$ appear twice in ch. This is rectified as follows: For each i in Y, find j in Y' and replace i with j in *ch′*, ensuring it adheres to the capacity constraint. Set $Y' = Y' \cap j$. Apply the same replacement procedure to address duplicates in *ch*. If these replacements prove infeasible due to capacity constraints, return to Step 1.

An example of the proposed crossover is shown in Figure 2. *C2* is selected from p and *C1* from p'. The lenght of the cut is 2. The cut in p is $X = 4, 5$ and the cut in p' is $X = 1, 4$. So, $Y = 5$ and $Y' = 1$ means that node 5 (resp. node 1) appears twice in p' (resp. in p) and must be replaced by 1 (resp. by 5).

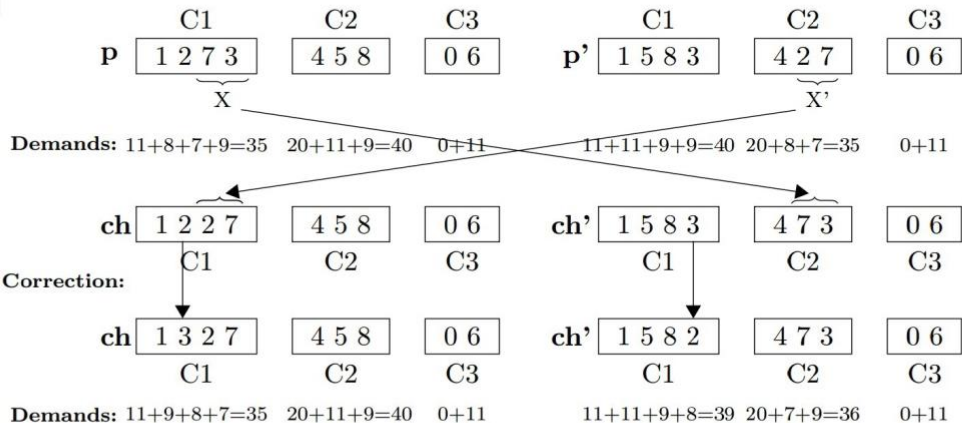


FIG .2 – An example of crossover

b. Mutation Operator

The purpose of the mutation operator is to broaden the exploration of the search space, mitigating the risk of getting trapped in local minima. Mutation is employed on 8% of the existing population and involves relocating a node from one cluster to another within the same solution. The target clusters for this relocation are those with the minimum number of nodes. It is crucial to note that, akin to the crossover operator, any alterations made to the clusters during mutation must undergo correction to ensure adherence to the specified constraints (the capacity constraint).

c. Re-schedule

While the crossover and mutation operators relocate nodes across clusters (routes), the reschedule operator serves the purpose of strategically reintegrating these nodes into the "optimal positions" within these routes. Suppose k is a node that has been moved to cluster C. In this scenario, the objective is to identify nodes u^* and v^* within C that minimize the expression $d(u, k) + d(k, v) - d(u, v)$ for all possible u and v in C. Subsequently, k is inserted between u^* and v^*. This process is iteratively applied to all clusters where nodes have been relocated.

4.2 Reinforcement Learning

The use of the Proximal Policy Optimization (PPO) algorithm [13], derived from Reinforcement Learning (RL), offers a promising approach to solving the VRP, a complex challenge involving optimal route planning for vehicles. The VRP requires an effective solution to satisfy demands while minimizing costs, and PPO stands out as a robust method to model and address this issue by enabling an agent to learn intelligent route policies.

In applying PPO to the VRP, the problem is formulated as an environment where an agent makes sequential decisions to maximize cumulative rewards. In the context of the VRP, the agent gains the ability to learn how to select optimal routes for vehicles, considering crucial parameters such as vehicle capacity and distances to be traveled. The PPO algorithm allows for a gradual adjustment of the agent's policy, ensuring adequate exploration of the solution space. The uniqueness of PPO lies in its mechanism of progressively updating policies, ensuring consistent and controlled improvements in complex decision-making. The underlying neural network of the algorithm takes an input state and generates the probability associated with each possible action in that state, guiding the agent to choose the most promising action. It is noteworthy that the number of neurons in the last layer of this network corresponds to the number of possible actions, establishing a direct correspondence between the network outputs and the decisions to be made. This neural approach provides PPO with the flexibility needed to handle various situations and adapt to the inherent complexity of the VRP.

The architecture proposed for our Proximal Policy Optimization algorithm typically includes a neural network to represent the agent's policy. Here is a general description of this architecture.

1. **Environment Observation (Input):** Observations of the environment, representing the current state of vehicles, serve as the input to the network.

2. **Neural Network (Policy):** A neural network is employed to represent the agent's policy. This network takes observations as input and generates a probability distribution over possible action.

3. **Action Sampling:** Actions are sampled from the probability distribution generated by the neural network, guiding the agent to make decisions based on the learned policy.

4. **PPO Loss Calculation:** The PPO loss is computed to update the neural network in a way that improves the agent's policy while limiting abrupt changes. The PPO loss is a function that measures the difference between the new and old policies, taking into account the advantage of actions [13].

The mathematical formula for the Proximal Policy Optimization (PPO) loss is as follows:

$$Loss\ (\theta) = E\ {*}\mathrm{min}\ (\frac{\rho_{\theta^{!}}(^{a}/_{s})}{\rho_{\theta}\ \ (^{a}/_{s})} A^{old}, clip(1-\epsilon, 1+\epsilon, \frac{\rho_{\theta^{!}}(^{a}/_{s})}{\rho_{\theta}\ \ (^{a}/_{s})}) A^{old} <=$$

Where:

- $Loss\ (\theta)$: PPO loss to be optimized
- $\rho_{\theta}\ \ (^{a}/_{s})$: Probability of selecting action a in state *s* according to the current policy θ.
- A^{old}: Advantage of the action calculated from the advantage function.
- ϵ : Clip hyperparameter to limit policy changes.

This mathematical formulation expresses the PPO loss, which is minimized through gradient descent during training to gradually improve the agent's policy while preventing drastic policy changes. The clipping mechanism limits the update size ensuring that the policy doesn't change too drastically.

5 Experimental results

This section is dedicated to the analysis of the obtained results. We highlight the effectiveness, efficiency, and practicality of each approach, thus providing a detailed evaluation of the performance of the techniques used. To compare GA and RL, we undertook to solve a real-world instance considered among the most complex. In this section, we analyze the evolution of the solution in terms of computation time and we compare the cost of solutions found by each algorithm. The following figure illustrates the progress of the cost reduction of the solution found by the GA as a function of the number of iterations.

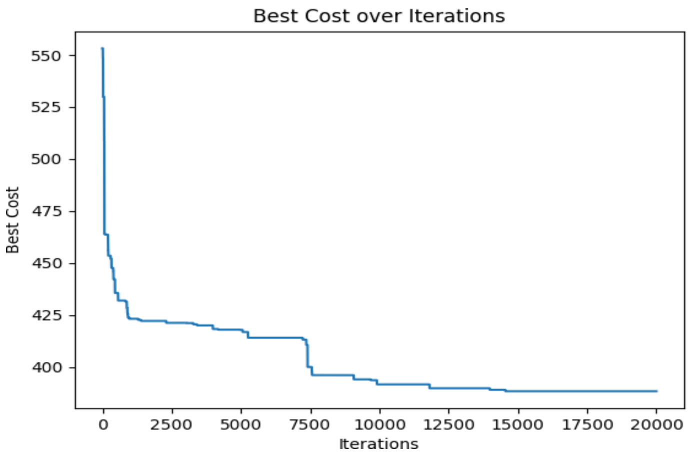


FIG .3 – The evolution of the cost of the solution using the GA

Figure 4 provides a graphical representation of the solutions obtained respectively by our GA and POO.

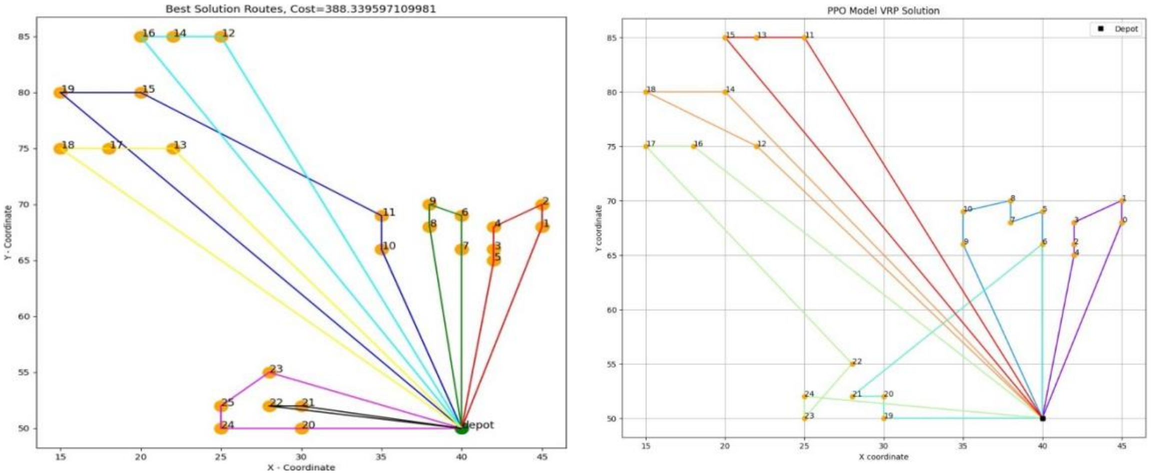


FIG .4 – Solutions provided by the GA vs RL.

The comparison of results obtained by the genetic algorithm and PPO to solve the vehicle routing problem reveals significant differences in terms of efficiency and computation time. While the genetic approach proves effective in searching for an optimized solution, it reached a cost of 388.34 after 20 000 iterations, indicating rapid but less precise convergence. In contrast, the PPO approach required a much deeper exploration with 1 million iterations to achieve a significantly lower cost of 285.4688, showcasing its ability to delve deeper into the search for more advanced optimization.

This increased efficiency of PPO can be particularly beneficial in contexts where cost reduction margins are critical, despite a heavier computational load and longer training time. This suggests that for large-scale problems where each unit of saved cost is paramount, PPO might be favored despite its resource and time requirements.

Conclusion

In conclusion, the utilization of both Genetic Algorithms (GA) and Reinforcement Learning (RL) in addressing the Vehicle Routing Problem (VRP) presents distinct advantages and trade-offs. The Genetic Algorithm, although efficient in quickly converging to a solution, may exhibit a less fine-tuned precision, as observed in the rapid convergence. This method proves effective for finding an optimized solution but may require careful consideration of convergence accuracy. On the other hand, Reinforcement Learning, specifically the Proximal Policy Optimization (PPO) algorithm, demonstrated a remarkable capability to explore the solution space thoroughly. While this approach demands a heavier computational load and longer training time, its increased efficiency becomes particularly advantageous in scenarios where cost reduction margins are crucial.

The decision to favor PPO over GA depends on the specific requirements of the problem at hand. In large-scale problems where each unit of saved cost holds paramount importance, the enhanced efficiency of PPO may outweigh its resource and time requirements. To address the latency issue of Proximal Policy Optimization (PPO), our perspective is to implement it in parallel and perform learning based on Graphics Processing Units (GPUs).

References

1. Ahmed, Z. H. and Yousefikhoshbakht, M. (2023). An improved tabu search algorithm for solving heterogeneous fixed fleet open vehicle routing problem with time windows. Alexandria Engineering Journal, 64, 349-363. https://doi.org/10.1016/j.aej.2022.09.008

2. Benaini, A., Berrajaa, A. and Daoudi, E. M. (2016). Solving the vehicle routing problem on gpu. In Proceedings of the Mediterranean Conference on Information and Communication Technologies 2015: MedCT 2015 Volume 2 (pp. 239-248). Springer International Publishing.

3. Berrajaa, A. and Benaini, A. (2022). GPU-based approach to large scale dynamic vehicle routing problem. International Journal of Logistics Systems and Management, 41(1-2), 225-242. https://doi.org/10.1504/IJLSM.2022.121006

4. Essaid, M., Idoumghar, L., Lepagnot, J. and Brévilliers, M. (2019). GPU parallelization strategies for metaheuristics: a survey. International Journal of Parallel, Emergent and Distributed Systems, 34(5), 497-522. https://doi.org/10.1080/17445760.2018.1428969

5. Joe, W. and Lau, H. C. (2020, June). Deep reinforcement learning approach to solve dynamic vehicle routing problem with stochastic customers. In Proceedings of the international conference on automated planning and scheduling (Vol. 30, pp. 394-402). https://doi.org/10.1609/icaps.v30i1.6685

6. Lin, B., Ghaddar, B. and Nathwani, J. (2021). Deep reinforcement learning for the electric vehicle routing problem with time windows. IEEE Transactions on Intelligent Transportation Systems, 23(8), 11528-11538. DOI: 10.1109/TITS.2021.3105232

7. Louati, A., Lahyani, R., Aldaej, A., Mellouli, R. and Nusir, M. (2021). Mixed integer linear programming models to solve a real-life vehicle routing problem with pickup and delivery. Applied Sciences, 11(20), 9551. https://doi.org/10.3390/app11209551

8. Mahmoudi, M. and Zhou, X. (2016). Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: A dynamic programming approach based on state space time network representations. Transportation Research Part B: Methodological, 89, 19-42. https://doi.org/10.1016/j.trb.2016.03.009

9. Reihaneh, M. and Ghoniem, A. (2019). A branch-and-price algorithm for a vehicle routing with demand allocation problem. European Journal of Operational Research, 272(2), 523-538. https://doi.org/10.1016/j.ejor.2018.06.049

10. Rouky, N., Abourraja, M., Boukachour, J., Boudebous, D., Alaoui, A. and Khoukhi, F. (2019). Simulation optimization based ant colony algorithm for the uncertain quay crane scheduling problem. International Journal of Industrial Engineering Computations, 10(1), 111-132. DOI: 10.5267/j.ijiec.2018.2.002

11. Rouky, N., Couzon, P., Boukachour, J., Boudebous, D., & Alaoui, A. E. H. (2018). Optimization of containers transfer in le havre port: A new algorithm for the railway transportation system. IFAC-PapersOnLine, 51(11), 1676-1681. https://doi.org/10.1016/j.ifacol.2018.08.215

12. Song, X., Jin, Y., Slabaugh, G. and Lucas, S. (2023). Joint action loss for proximal policy optimization. arXiv preprint arXiv:2301.10919. https://doi.org/10.48550/arXiv.2301.10919

13. Xue, D., Wu, D., Yamashita, A. S. and Li, Z. (2023). Proximal policy optimization with reciprocal velocity obstacle based collision avoidance path planning for multi- unmanned surface vehicles. Ocean Engineering, 273, 114005. https://doi.org/10.1016/j.oceaneng.2023.114005

14. Abourraja, M. N., Benantar, A., Rouky, N., Boudebous, D., Boukachour, J., & Duvallet, C. Towards a simulation-based decision support tool for container terminal layout design (2019) International Conference on Harbour, Maritime and Multimodal Logistics Modelling and Simulation, pp. 11-17. DOI: 10.46354/i3m.2019.hms.002

Comparative Analysis of Metaheuristic Algorithms for Vehicle Routing Problem

Manal El Jaouhari[1], Ghita Bencheikh[2], Ghizlane Bencheikh[3]

[1] Moulay Ismaïl University of Meknes, Morocco
man.eljaouhari@edu.umi.ac.ma
[2] LINEACT, Cesi Engineering School, Assat, France
gbencheikh@cesi.fr
[3] Moulay Ismaïl University of Meknes, Morocco
g.bencheikh@umi.ac.ma

Abstract : Optimization is a constantly evolving field, with new methods emerging to tackle complex problems. Metaheuristic algorithms, known for their flexibility in solving difficult optimization tasks, have become increasingly popular. In this study, we compare four metaheuristics: Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Hill Climbing (HC), and Genetic Algorithm (GA) against exact methods, to solve combinatorial optimization problems, with a focus on the Capacitated Vehicle Routing Problem (CVRP) as a test case. Our goal is to identify the strengths and weaknesses of each algorithm in terms of solution quality and computational efficiency. The results highlight ACO's accuracy, PSO's balanced approach, HC's speed, and GA's relative inefficiency providing insights for future multi-agent system approaches.

Keywords : *Metaheuristic, Hill Climbing, Ant Colony Optimization, Particle Swarm Optimization, Genetic Algorithm, Exact Algorithm, Comparative Study.*

1 Introduction

A combinatorial optimization problem involves finding an optimal solution to an objective function defined within a search space, either maximizing or minimizing the function under a set of constraints specific to the problem. The possible outcomes in solving these problems are classified as optimal solutions, near-optimal solutions, or feasible solutions. To address such problems, a wide variety of algorithms have been developed and utilized in the literature, generally categorized into two main groups: exact algorithms and approximate algorithms.

1. **Exact Algorithms:** These algorithms guarantee an exact solution and are typically employed for smaller problems. However, as the complexity and size of the problem grow, exact algorithms become computationally intensive, requiring an enormous amount of time, which makes solving large or complex problems highly challenging.

2. **Approximate Algorithms:** These algorithms, on the other hand, provide high-quality solutions within a reasonable computational time. Metaheuristics, a subset of approximate algorithms, are particularly useful for complex problems. They are preferable because they can efficiently yield satisfactory results without the excessive computational demands associated with exact methods.

This study conducts a comparative analysis of various metaheuristic algorithms for optimization, using the Vehicle Routing Problem (VRP) as a case study. The VRP is one of the most extensively researched topics in operations research and combinatorial optimization, largely due to its critical role in real-world logistics and distribution systems, which are essential for the smooth functioning of modern economies. Given the multitude of logistical constraints and diverse scenarios encountered in practice, the VRP has evolved into numerous variants that address different operational challenges. These variants include time windows, multiple depots, mixed vehicle fleets, split deliveries, pickups and deliveries, precedence constraints, and complex loading conditions. Among the variants, the Capacitated Vehicle Routing Problem (CVRP) stands out as the fundamental problem. First introduced by Dantzig and Ramser in 1959 [4]. The CVRP aims to optimize the routing of a fleet of vehicles with fixed carrying capacities to service a set of customers with known demands. The goal is to minimize transportation costs while ensuring that each vehicle's capacity is not exceeded and all customers are served from a central depot. Over the years, the CVRP has been adapted to address real-world complexities such as uncertain demand, multiple service levels, and environmental factors, cementing its importance in logistics and transportation.

This analysis focuses on four metaheuristic approaches to solving the VRP:

- **Hill Climbing (HC)**, which improves an initial solution through local search;
- **Ant Colony Optimization (ACO)**, inspired by the foraging behavior of ants to construct efficient routes;
- **Particle Swarm Optimization (PSO)**, where solutions are modeled as particles that adjust based on personal and group experiences;
- **Genetic Algorithm (GA)**, which evolves solutions using genetic operations like selection, crossover, and mutation.

We have implemented four metaheuristics, and by comparing these algorithms in terms of solution quality and computational efficiency, this study aims to elucidate their respective strengths and weaknesses in addressing the optimization challenges posed by the CVRP. Through this comparative analysis, we assess the ability of each metaheuristic to generate near-optimal solutions while considering the trade-offs between accuracy and computational cost. By doing so, we aim to provide insights into the most suitable metaheuristics for various instances of the CVRP, depending on the problem's complexity and computational constraints.

The structure of the paper is organised as follows: Section 2 provides a comprehensive overview of the related studies in the field. In Section 3, we present the mathematical formulation of the CVRP . Section 4 details the four metaheuristics utilized in this study: Hill Climbing 4.1, Ant Colony Optimization 4.2, Genetic Algorithm 4.3, and Particle Swarm Optimization 4.4, along with their corresponding algorithms. In Section 5, we summarize the experimental results in a table 1, which offers a comparative analysis of the performance of each algorithm in terms of solution quality and computational efficiency. This comparison facilitates a thorough evaluation of the strengths and limitations of each metaheuristic. Subsequently, Section 6 discusses the analysis of the results, and the paper concludes in Section 7 with final insights and directions for future research.

2 Insights from Recent Studies

In vehicle routing optimization, metaheuristic algorithms have been widely adopted to tackle complex logistical challenges. Our study builds on this foundation by comparing the performance of four metaheuristics in solving the CVRP, aiming to evaluate their strengths and limitations in terms of solution quality and computational efficiency. Several key studies contribute to the state of the art in this domain. Tiwari et al. [13] explored various local search algorithms, highlighting the strengths of Tabu Search for larger instances of vehicle routing problems, which aligns with our interest in evaluating computational efficiency across different problem scales. Tadaros et al. [12] tackled a complex variant of the VRP in Nordic logistics using General Variable Neighborhood Search (GVNS), providing a benchmark for comparing its performance against the metaheuristics we investigate. Shi et al. [11] advanced the field by integrating Adaptive Large Neighborhood Search (ALNS) with zone-based pricing strategies, showing its effectiveness in optimizing both routing and pricing decisions, which informs our exploration of adaptive approaches. Furthermore, the use of Harris Hawks Optimization (HHO) by Xue et al. [14] demonstrates superior performance in minimizing objective functions for VRP, an aspect we examine with ACO and PSO. Studies such as Shi et al. [10], which incorporate real-time capabilities for UAV-based VRP solutions, and Matijevic et al. [7], which apply metaheuristics to real-world transportation logistics, further illustrate the versatility of these algorithms in diverse contexts. Despite the diversity of approaches, most studies focus on specific algorithms or settings, limiting comparisons across different contexts, While metaheuristics are increasingly important in solving vehicle routing problems, there is a lack of comparative studies evaluating multiple algorithms across various instances of the same problem. Our study compares the performance of the four metaheuristics, HC, ACO, PSO and GA, within the framework of the CVRP as test case. The goal is to measure their relative performance in terms of solution quality and computational efficiency, providing a comprehensive comparison that highlights their strengths and limitations across different problem instances.

3 Mathematical Model

Notation

N: Set of nodes, including the depot and customers.

K: Number of vehicles.

d_{ij}: Euclidean distance between node i and node j.

q_j: Demand of customer j.

C: Vehicle capacity.

x_{ijk}: Binary decision variable that equals 1 if vehicle k travels from node i to node j, and 0 otherwise.

u_{ik}: Continuous variable representing the load of vehicle k when leaving node i.

Objective Function

The objective is to minimize the total distance traveled by all vehicles:

$$Min\, f(x) = \sum_{k=1}^{K} \sum_{i \in N} \sum_{j \in N} d_{ij} \cdot x_{ijk} \tag{1}$$

Constraints

1. **Capacity Constraints**: Ensure that the total demand served by each vehicle does not exceed its capacity:

$$\sum_{j \in N} q_j \cdot \sum_{i \in N} x_{ijk} \leq C, \quad \forall k \in K \tag{2}$$

2. **Visit Constraints**: Each customer must be visited exactly once by one vehicle:

$$\sum_{i \in N} \sum_{k=1}^{K} x_{ijk} = 1, \quad \forall j \in N \tag{3}$$

3. **Flow Conservation Constraints**: Ensure that if a vehicle arrives at a node, it must leave that node:

$$\sum_{j\in N} x_{ijk} = \sum_{j\in N} x_{jik}, \quad \forall i \in N, \forall k \in K \tag{4}$$

4. **Depot Constraints**: Each vehicle must start and end at the depot:

$$\sum_{j\in N} x_{0jk} = 1, \quad \forall k \in K \tag{5}$$

$$\sum_{j\in N} x_{j0k} = 1, \quad \forall k \in K \tag{6}$$

5. **Subtour Elimination (MTZ Constraints)**: The Miller-Tucker-Zemlin formulation: Proposed by Miller in 1960 [9]. It was initially proposed for a vehicle routing problem (VRP), where each route's number of vertices is limited [15]. It prevents subtours construction by using vehicle loads variables:

$$u_{ik} - u_{jk} + (n-1)\cdot x_{ijk} \le n-2, \quad \forall i, j \in N, \forall k \in K \tag{7}$$

4 Metaheuristic Algorithms

Metaheuristic algorithms can be broadly classified into two categories: single-solution-based algorithms and population-based algorithms [1]. In our research, we have tested algorithms from both classifications.

↬ For single-solution-based metaheuristics, we evaluated Hill Climbing (HC) and Ant Colony Optimization (ACO).

↬ For population-based algorithms, we explored Genetic Agorithm (GA) and Particle Swarm Optimization (PSO).

4.1 Hill Climbing

Hill Climbing (HC) is a straightforward yet powerful local search algorithm that iteratively refines a single solution by exploring its neighboring alternatives within the search space [6]. Starting from an initial feasible solution, HC seeks to improve upon it by transitioning to a neighboring solution, but only if it offers a better objective value. The name "hill climbing" is derived from the fact that, in a fitness landscape with a maximization aim, the local maximum is a hill, where the individual must always walk upwards, eventually reaching the peak. The initial individual is randomly generated, and mutation is applied to generate new candidates. If the candidate's solution is weaker, it is rejected; if it is better, the individual moves up to it. This process continues until no further improvements can be made, typically resulting in convergence at a local optimum. While Hill Climbing is computationally efficient and easy to implement, its major limitation lies in its tendency to become trapped in local optima, preventing it from reaching a global solution. In the context of the CVRP, HC can be applied by first generating an initial set of routes for vehicles, each assigned a group of customers while adhering to capacity constraints. The algorithm then refines these routes by making localized adjustments, such as exchanging customers between routes or reordering delivery sequences, to minimize the overall travel distance. This iterative process continues until the route configuration reaches a point where no further improvements can be achieved, yielding a locally optimized solution.

4.1.1 HC Algorithm for CVRP

The following pseudo-code 1 showcases its iterative process of refining an initial solution by exploring neighboring solutions. The algorithm begins by evaluating the quality of the current solution x using the objectify function f (1). The best neighbor of the current solution is selected and the current solution is updated. This process of exploring and updating continues until no better neighboring solution is found. At that point, the algorithm terminates and returns the best solution discovered. While the algorithm is designed to efficiently find an optimal or near-optimal solution, it may stop prematurely at a local optimum due to its greedy strategy of only accepting improvements.

Algorithm 1 Hill Climbing Algorithm

```
Input:
    $G(V, E, d)$: graph of customers to visit,
    $C$: Capacity of vehicles,
    $x_0$: Initial solution,
Variables:
    $x : Currentsolution$
Output:
    $x* : Bestsolution(besttour)$

$x \leftarrow x_0$;
$x* \leftarrow x_0$;
while True do
    $x \leftarrow \arg\min(f(x_i)) \qquad \forall x_i \in V(x)$;
    if $f(x) < f(x*)$ then
        $x* \leftarrow x$;
    else
        Break;
    end if
end while
```

4.2 Ant Colony Optimization

Ant Colony Optimization (ACO), introduced by Dorigo et al. [2], is a highly influential metaheuristic inspired by the foraging behavior of real ants. ACO is particularly well-suited for solving combinatorial optimization problems, such as routing, where the search space is vast and complex. The algorithm mimics the way real ants deposit pheromones to mark favorable paths as they search for food. In the artificial setting, a colony of artificial ants builds solutions by following pheromone trails left by previous ants, which are combined with problem-specific heuristics, such as customer proximity or demand in routing problems. As ants traverse different paths, successful solutions are reinforced through the deposition of more pheromones on those paths, thereby increasing the likelihood that future ants will explore these promising routes. This dynamic process allows ACO to efficiently navigate large solution spaces while balancing exploration and exploitation [3].

When applied to the CVRP, ACO simulates the process of ants constructing vehicle routes. Each ant builds a complete solution by visiting a sequence of customers, with their decisions influenced by pheromone levels and distance-based heuristics that favor shorter and more efficient paths. After all ants have constructed their routes, the most successful solutions—those minimizing total travel distance while respecting vehicle capacity constraints—are reinforced by higher pheromone levels, encouraging subsequent ants to focus on these efficient paths. This iterative process ensures that all customers are served within vehicle capacity limits, and through the collective behavior of the artificial ants, ACO converges toward optimal or near-optimal routing solutions for the CVRP.

4.2.1 ACO Algorithm for CVRP

The following pseudo-code 2 operates iteratively by simulating the behavior of ants that build routes for the CVRP based on pheromone trails and heuristic information. The algorithm begins by initializing the pheromone matrix τ with default values and creating a population of ants for each iteration. Each ant starts at the depot and sequentially visits customers by calculating the probabilities of choosing the next customer based on pheromone levels and distance-based heuristics (controlled by the parameters α and β). If an ant's vehicle becomes full or the next node is the depot, the ant returns to the depot to complete its route. Once all ants have constructed their routes, the pheromone matrix is updated to reinforce successful routes while allowing pheromone evaporation to promote exploration. The best solution from each iteration is compared to the global best solution, and if it is better, the global best is updated. After completing the predefined number of iterations, the algorithm returns the best solution found and its cost. The primary goal is to minimize the total travel distance for the CVRP while ensuring all customers are serviced within the vehicle's capacity.

4.3 Genetic Algorithm

The Genetic Algorithm (GA) is developed by John Holland and his students [5] between 1960 and 1970. The GA is a population-based metaheuristic inspired by biological evolution, including mechanisms such as natural selection, crossover, and mutation [6]. A population of candidate solutions evolves over generations, where fitter individuals are selected to produce offspring through crossover, while mutation introduces random changes to maintain diversity and prevent premature convergence. This process, inspired by Darwin's theory of survival of the fittest, allows GA to explore a wide solution space and converge toward optimal or near-optimal solutions. The population pool undergoes modifications through genetic operators, with selection determining which individuals are used for crossover and mutation. In problems like the CVRP, GA begins by generating an initial population of vehicle routing solutions. Over generations, the algorithm refines these solutions by selecting the most efficient routes, combining them through crossover, and occasionally introducing mutations to improve the routes further. The objective is to minimize total travel distance while respecting vehicle capacity constraints, with the process continuing through iterations until a stopping criterion is met, ultimately converging to a solution that balances efficiency and feasibility.

4.3.1 GA Algorithm for CVRP

This pseudo-code implements the GA, a population-based optimization technique inspired by biological evolution. The algorithm begins by initializing a population of candidate solutions, which represent possible vehicle routes. The fitness of each solution is evaluated using the objectif function f (eq 1, called fitness function. In each generation, the algorithm selects two parents from the population based on their fitness value. These parents are then combined using the Crossover() function to produce offspring. A mutation is applied to the offspring with a certain probability using the Mutate() function to introduce diversity and prevent premature convergence. The new population is formed by replacing the selected parents by the new generated individual after offspring operations. This process continues over several generations. The algorithm tracks the best solution encountered during the search.

4.4 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population-based optimization algorithm inspired by the collective behavior of birds or fish schools [8]. Each particle represents a potential solution that moves through the search space, adjusting its position based on its own best-known solution (personal best) and the best solutions found by other particles (global best). The particles explore the search space by learning from their personal experiences and the best solutions discovered by the swarm, converging towards optimal or near-optimal solutions over

Algorithm 2 Ant Colony Optimization algorithm

```
  Input:
    $G(V, E, d)$: graph of customers to visit
    $C$: Capacity of vehicles,
    $n$: number of ants,
    $H$: maximum number of iterations,
    $\eta_{ij} = \frac{1}{d_{ij}}$ is a heuristic value,
    $\tau_{ij}$: the pheromone concentration,
    $\mu_{ij} = d_{i0} + d_{0j} - d_{ij}$ is the savings of combining two cities $i$ and $j$ on one
      tour as opposed to visiting them on two different tours,
    $\alpha, \beta, \gamma$: the relative influence of the pheromone concentration, the heuristic value,
      and the savings values
    $\epsilon$: evaporation rate,
    $ants$: set of ants,
  Variables:
    $\bar{V}$: Remaining customers (not visited),
    $i$: Current node,
    $j$: Next node,
    $L_k$: current load,
    $P_{ij}^k$: probability distribution,
    $\bar{x}$: Current best solution (best tour)

  Output:
    $x* : Bestsolution(besttour)$

  Initialize $\tau$ with default values
  for $it \in \{1, \dots, H\}$ do
      for each k ∈ ants do
          $i \leftarrow 0$
          $\bar{V} \leftarrow V$
          $L_k \leftarrow 0$
          while $\bar{V} \neq \emptyset$ do
              $P_{ij}^k \leftarrow \frac{\tau_{ij}^\alpha \eta_{ij}^\beta \mu_{ij}^\gamma}{\sum_{l \in \bar{V}} \tau_{ij}^\alpha \eta_{ij}^\beta \mu_{ij}^\gamma} \forall j \in \bar{V}$
              $j \leftarrow \arg\max(P)$
              if $j = 0 || L_k \geq C$ then
                  go back to the depot
                  $L_k \leftarrow 0$
              else
                  Visit customer $j$
                  $L_k \leftarrow d_j$
                  $\bar{V} \leftarrow \bar{V} \setminus \{j\}$
              end if
          end while
          go back to the depot
      end for
      Update Pheromones
      $\bar{x} \leftarrow \arg\min(f(k)) \qquad \forall k \in ants$
      if $f(\bar{x}) < f(x*)$ then
          $x^* \leftarrow \bar{x}$
      end if
  end for
```

Algorithm 3 Genetic Algorithm Algorithm

```
Input:
    $G(V, E, d)$: graph of customers to visit
    $C$: Capacity of vehicles,
    $n_p$: population size,
    $\tau_m$: mutation rate,
    $\tau_c$: crossover rate,
    $H$: maximum number of iterations,
Variables:
    $P$: Set of individuals (solutions),
    $x*$: Best solution (best tour),
Output:
    $x*$: Best solution (best tour),

$P_i \leftarrow$ random solution $\quad \forall i \in \{1, ..., n_p\}$
for $it \in \{1, \ldots, H\}$ do
    $parent1 \leftarrow random(x \in P)$
    $parent2 \leftarrow random(x \in P)$
    if $Random() < \tau_c$ then
        $parent1, parent1 \leftarrow Crossover(parent1, parent2)$
    end if
    if $Random() < \tau_m$ then
        $parent1 \leftarrow Mutate(parent1)$
    end if
    if $Random() < \tau_m$ then
        $parent2 \leftarrow Mutate(parent2)$
    end if
end for
$x* \leftarrow \arg\min(f(k)) \qquad \forall k \in P$;
```

time [6]. In the context of the CVRP, each particle represents a set of vehicle routes. Particles explore different routing configurations, refining their solutions by learning from their personal experiences and the global best found by the swarm. The algorithm's velocity update is influenced by three key strategies: moving towards personal best, moving towards global best, and incorporating momentum to maintain direction. The goal of PSO in CVRP is to minimize total travel distance while respecting vehicle capacity constraints, with particles continuously exploring the search space and avoiding premature convergence on local optima.

4.4.1 PSO Algorithm for CVRP

The following pseudo-code 4 implements the PSO algorithm, a population-based optimization technique inspired by the social behavior of birds flocking or fish schooling. The algorithm begins by initializing a swarm of particles, where each particle represents a potential solution to the CVRP. Each particle's position and velocity are set randomly, and the initial solution is evaluated to compute the particle's personal best position. The global best solution, found among all particles, is also identified at the start. In each iteration, the algorithm evaluates the cost of the current particle's position. If the cost is better than the current global best, the global best position and cost are updated. Each particle also updates its personal best position if its current cost is lower than its previous best. The algorithm then updates the velocity of each particle according to the PSO update rule, which considers both the particle's personal best and the global best solution. The new velocity is used to adjust the particle's position, and the resulting solution is evaluated. If the new solution improves the particle's personal best, it is updated accordingly. Similarly, if any particle discovers a better solution than the global best, the global best is updated. This iterative process continues for a set number of iterations, and at the end of the algorithm, the best solution found is returned as the optimal solution. The primary objective of the PSO algorithm in the CVRP is to optimize vehicle routes by minimizing the total travel distance while ensuring that vehicle capacity constraints are respected, as specified in the input parameters (e.g., number of customers, vehicles, and vehicle capacities).

Algorithm 4 Particle Swarm Optimization Algorithm

```
Input:
  $G(V, E, d)$: graph of customers to visit
  $C$: Capacity of vehicles,
  $n_p$: number of particles,
  $H$: maximum number of iterations,
Variables:
  $P$: Set of particles,
  $v$: Velocity of particles,
Output:
  $x*$: Best solution

Initialize particles with random positions and velocities
for $it \in \{1, \ldots, H\}$ do
    for each $p \in P$ do
        if f($p$) < f(x*) then
            $x* \leftarrow p$
        end if
        $p \leftarrow Updatepersonalbestposition$
    end for
    for each $i \in P$ do
        Update velocity for particle $i$:
            $v_i \leftarrow w \times v_i + c_1 \cdot r_1 \cdot (p_{best,i} - x_i) + c_2 \cdot r_2 \cdot (g_{best} - x_i)$
        Update position for particle $i$:
            $x_i \leftarrow x_i + v_i$
    end for
end for
```

5 Evaluation of the Metaheuristic Algorithms

The methodology involves solving the CVRP using four metaheuristic algorithms: HC, PSO, GA, and SA, all implemented in Python. Random instances of the CVRP are generated, where vehicles with limited capacity must meet customer demands while minimizing travel distance. The exact solution is obtained using the GLPK solver in Julia. Each algorithm is applied to these instances, and their performance is evaluated by comparing solution quality and computational efficiency across randomly generated benchmark instances. The results of each algorithm are compared to the exact solution by calculating the gap, which measures the percentage difference between the metaheuristic solutions and the optimal solution. This comparative analysis aims to identify the strengths and weaknesses of each metaheuristic in solving the CVRP.

The performance Gap for a given metaheuristic compared to the exact method is calculated using eq 8.

$$\text{Gap (\%)} = \left(\frac{\text{BD}_{\text{Metaheuristic}} - \text{BD}_{\text{Exact}}}{\text{BD}_{\text{Exact}}} \right) \times 100 \tag{8}$$

5.1 Experimental Setup

All experiments were conducted using random instances. The computational platform employed for this study was a laptop equipped with an Intel(R) Core(TM) i5-8300H CPU running at 2.30 GHz and 8 GB of RAM from the 8th generation. This hardware configuration was chosen to showcase the practical applicability and efficiency of the algorithms for the CVRP and to compare the results with the solution found by the exact method.

5.2 Key Features and Parameters

The parameters used in the algorithms are carefully selected to optimize their performance.

- For the ACO, the number of ants is set to 100, with a pheromone importance ($\alpha = 1$), heuristic information importance ($\beta = 3$), an $\epsilon = 2$, and a pheromone deposit factor $\tau = 70$.
- For the GA, the population size is 150, with 500 generations and a mutation rate of 0.1.
- In PSO, the algorithm is configured with 50 particles, 200 iterations, an inertia weight of 0.4, a cognitive weight of 1.5, and a social weight of 1.7.

These parameters are selected to balance exploration and exploitation by ensuring sufficient diversity in the search process (e.g., through larger populations and multiple iterations) while guiding the algorithm toward promising solutions (e.g., through pheromone updates, social influences, and mutation).

6 Analysis

The table presents the performance of the four metaheuristic algorithms—ACO, HC PSO, and GA—applied to randomly generated instances of the CVRP, with comparisons to exact methods. We chose to test a manageable set of cities (5, 10, and 15) and vehicles (2, 3, and 5), generating 10 random instances for each combination. This range was selected to strike a balance between problem complexity and computational feasibility.

In the Exact (BD - Best Distance) section, the best-known solutions from the exact method serve as lower bounds for the metaheuristics. The best solution found by each algorithm is compared to this exact solution to assess the quality of the results. The Gap (%) column shows the percentage difference between the metaheuristic solutions and the optimal solution. A lower gap percentage indicates a better approximation of the optimal solution. Finally, the Computation Time (CPU Time) column displays the time taken by each method to compute the solution, helping evaluate the trade-off between solution quality and computational efficiency.

TAB. 1: Comparison of the Performance of the Algorithms

Instance	Exact		ACO			PSO			HC			GA		
	BD	CPU(s)	BD	Gap (%)	CPU(s)	BD	Gap (%)	CPU(s)	BD	Gap (%)	CPU(s)	BD	Gap (%)	CPU(s)
X-n5-k2	632.6460	0.0013	737.3459	16,5495	0.6423	737.3459	16,5495	0.0015	737.3459	16,5495	3.4149	730.1117	15,4060	0.2787
A-n5-k2	648.2130	3.2918	1358.0071	109,5001	0.8815	1358.0071	109,5001	0.0240	1082.9749	67,0708	5.0173	762.1134	17,5714	0.2273
B-n5-k2	568.3043	0.0007	920.6234	61,9947	0.7091	920.6234	61,9947	0.0016	1073.5494	88,9040	11.7238	856.5037	50,7122	0.8327
C-n5-k2	759.9770	0.0010	1353.6225	78,1136	0.6247	1991.9993	162,1131	0.0019	1205.6581	58,6440	5.0043	969.4958	27,5691	0.1912
D-n5-k2	492.6919	0.0009	793.4884	61,0516	0.6936	1253.1543	154,3485	0.0024	550.5554	11,7444	3.2198	778.1037	57,9291	0.2961
E-n5-k2	891.4364	0.0006	1054.6053	18,3040	0.7196	1054.6053	18,3040	0.0017	1054.6053	18,3040	11.2251	1049.5831	17,7407	0.3449
F-n5-k2	773.1259	3.5409	1275.9087	65,0325	0.6569	1430.1159	84,9784	0.0013	1126.0133	45,6442	4.6722	958.2789	23,9486	0.1999
G-n5-k2	956.7626	0.0010	1205.6582	26,0144	0.6786	1205.6582	26,0144	0.0018	1205.6581	26,0143	3.5078	959.3112	0,2664	0.2905
H-n5-k2	597.1392	0.0007	829.6064	38,9302	0.6306	829.6064	38,9302	0.0030	829.6064	38,9302	3.6645	658.1390	10,2153	0.3494
I-n5-k2	742.2421	0.0007	931.5999	25,5116	0.6852	931.5999	25,5116	0.0014	931.5998	25,5116	3.7188	916.0347	23,4145	0.3131
X-n10-k3	1027.6019	0.0237	1105.3965	7,5705	5.7433	1105.3965	7,5705	0.0207	1548.7314	50,7132	7.8790	1060.8636	3,2368	0.3629
A-n10-k3	1202.9948	0.1872	1445.9922	20,1994	5.1131	1591.3207	32,2799	0.0170	2160.3078	79,5775	18.4668	1736.7614	44,3698	0.8237
B-n10-k3	1385.3676	0.0652	1488.9701	7,4783	5.4658	1590.4550	14,8038	0.0266	1747.5083	26,1404	7.6644	1698.1143	22,5750	0.9505
C-n10-k3	978.9489	0.2502	1043.2374	6,5671	1.9157	1184.3220	20,9789	0.0165	1284.0174	31,1629	8.3398	1076.0501	9,9189	0.4345
D-n10-k3	1273.1859	0.0316	1345.6565	5,6921	2.1337	1410.9831	10,8230	0.0130	1618.8182	27,1470	19.3497	1407.0059	10,5106	0.3069
E-n10-k3	1159.7094	0.1352	1344.8954	15,9683	1.4901	1344.8954	15,9683	0.0207	1931.4923	66,5497	18.3500	1561.5705	34,6519	0.3965
F-n10-k3	967.2625	0.0443	1205.1432	24,5932	1.9442	1205.1432	24,5932	0.0202	1578.3889	63,1810	17.0545	1312.8234	35,7257	0.3379
G-n10-k3	982.6145	0.1223	1119.7954	13,9608	1.9904	1219.6917	24,1272	0.0085	1216.5738	23,8099	19.7098	1226.3541	24,8052	0.3473
H-n10-k3	1009.3241	0.0647	1164.7702	15,4010	2.3681	1164.7702	15,4010	0.0368	1511.8412	49,7875	7.8897	1506.6458	49,2727	0.3494
I-n10-k3	1813.6170	3.8154	2010.8946	10,8776	2.3819	2012.1701	10,9479	0.0062	2022.5392	11,5196	19.0502	1927.6304	6,2865	0.3831
X-n15-k5	1211.3604	0.4246	1292.7975	6,7228	4.5217	1292.7975	6,7228	0.1402	2459.6503	103,0486	13.9031	1814.1539	49,7617	0.2681
A-n15-k5	1669.7577	1.2091	1903.1864	13,9798	4.6785	2255.2393	35,0639	0.0693	2396.3247	43,5133	21.1818	2346.5249	40,5636	0.8475
B-n15-k5	1354,0549	78,3214	1797.4811	10,5877	3.6573	1797.4811	10,5877	0.0614	1886.5209	16,0382	18.0605	1727.9580	6,2935	0.3982
C-n15-k5	1318,8295	43,763	1784.2213	13,3795	3.7599	1962.1439	24,6312	0.0519	1976.2048	25,5923	18.4110	1817.6743	15,4851	0.3862
D-n15-k5	1347,0481	0,3087	1493.6579	11,8327	4.2318	1637.3254	22,5983	0.0192	1764.8257	32,1579	12.5234	1581.9941	16,0325	0.3371
E-n15-k5	1380,8569	1,961	1814.5049	12,7162	2.9844	1896.2761	17,7937	0.0205	2021.1435	25,5073	17.7385	1861.6574	14,5061	0.4043
F-n15-k5	1618,8564	0,0843	1698.2317	10,4159	3.3415	1773.9624	15,3574	0.0631	1874.0364	22,1412	17.0143	1769.1211	15,0181	0.2938
G-n15-k5	1440,451	4,7682	1857.3945	7,2742	3.1428	2034.8799	17,5187	0.0321	2065.2382	19,2593	18.3434	1994.6397	13,1944	0.4219
H-n15-k5	1356,5448	1,1349	1622.8744	15,0924	2.5375	1622.8744	15,0924	0.0247	1812.4538	28,5237	14.6734	1697.0978	20,3356	0.3249
I-n15-k5	1565,3149	8,7458	1713.2408	14,6962	2.7618	1825.6005	21,7908	0.0398	1975.3452	32,2475	17.6598	1862.4498	24,6533	0.395

Based on the analysis of gaps (see Figure 2) and CPU times (see Figure 1) for the different metaheuristic algorithms, several key observations can be made. For the ACO algorithm, the results consistently show the lowest mean gap values across most instances, meaning it produces solutions that are closer to the optimal solution compared to other algorithms. For example, when tested on the instance with n=10 and k=3, ACO had a mean gap of 12.83%, with a standard deviation of 6.33%, reflecting that its solutions are both relatively consistent and accurate. However, ACO comes at the cost of higher computational time, taking an average of 3.05 seconds on the same instance. This suggests that while ACO is the most accurate, it requires more time to arrive at these solutions.

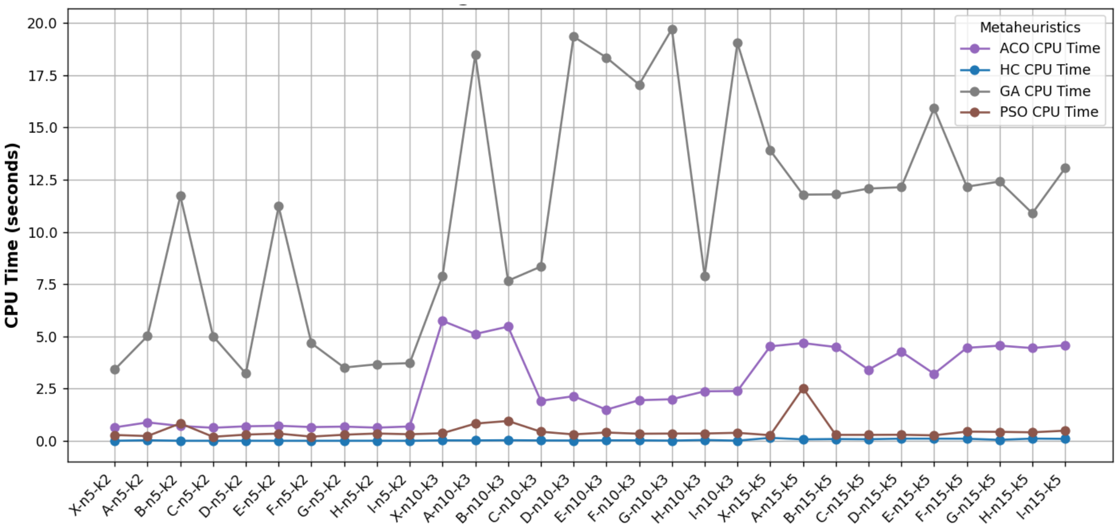


FIG. 1: CPU Time (in seconds) for ACO, HC, GA and PSO across different problem instances

In contrast, HC demonstrated significantly faster computational times. For the $n = 10$, $k = 3$ instance, the mean CPU time for HC was only 0.02 seconds, which is much faster than ACO. However, its mean gap was 17.75%, indicating that the solutions it finds are not as close to the optimal solution. This highlights the trade-off between speed and solution quality, with HC being much faster but less accurate.

The GA exhibited the highest gaps and relatively longer computation times. For n=10, k=3, the GA had a mean gap of 42.96 % and took an average of 14.38 seconds to compute the solution. This suggests that GA struggles both in solution quality and efficiency, making it the least favorable option overall.

Lastly, PSO provided a balanced approach. While it did not achieve the same level of accuracy as ACO, it performed reasonably well in both solution quality and speed. For the n=10, k=3 instance, the mean gap for PSO was 24.14%, and the mean CPU time was 0.47 seconds. This shows that PSO strikes a good compromise, offering a decent solution quality while remaining faster than both ACO and GA.

To further illustrate the performance of the algorithms, Figure 2 presents a boxplot comparing the gap values of the metaheuristics to the exact solution. The boxplot visualizes the variability and central tendency of the gap values for each algorithm. ACO consistently achieves the smallest gaps, confirming its status as the most accurate algorithm. In contrast, GA shows the largest gaps and the highest variability, indicating that it frequently struggles to approximate the optimal solution. In addition to the box, individual outliers are marked by circles. These outliers represent data points that fall significantly outside the general range, indicating poor performance or unusually good results for specific instances. For example, GA shows several outliers in the higher gap range, suggesting occasional instances where it performed poorly in comparison to other algorithms. The box and its associated statistics clearly compare the algorithms' consistency and effectiveness across different problem instances. This visual representation reinforces the numerical findings, showing that ACO delivers superior accuracy,

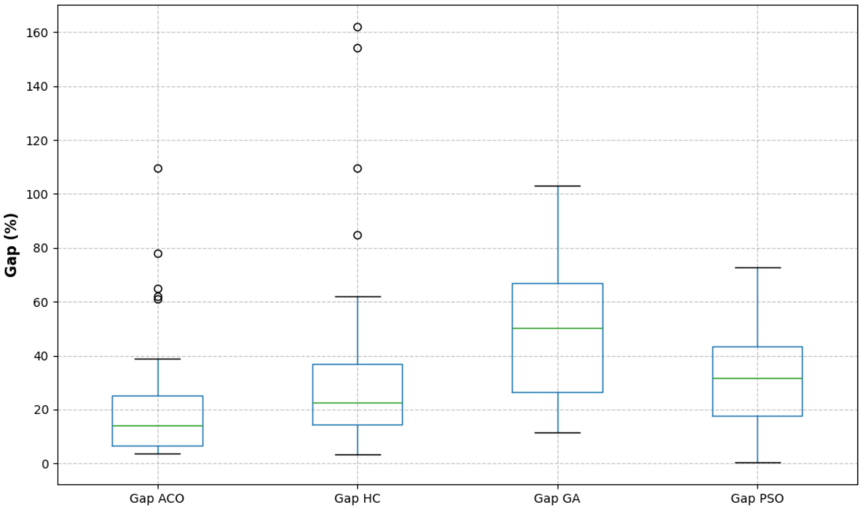


FIG. 2: Gaps of metaheuristics to the exact solution

while GA exhibits larger gaps and higher variability. PSO offers a reasonable compromise, balancing solution quality and computational time, while HC emphasizes computational speed over solution accuracy.

7 Conclusion and perspectives

In this study, we investigated the performance of four metaheuristics -ACO, PSO, HC, and GA- for solving the CVRP using randomly generated instances. The primary objective was to conduct a comprehensive comparison of these algorithms, evaluating their strengths and limitations in terms of solution quality and computational efficiency. After implementing and benchmarking each metaheuristic, the gap between their solutions and the exact solution was computed.

The results demonstrate distinct characteristics for each algorithm: ACO exhibits high accuracy, though at the cost of longer execution times; PSO achieves a balanced trade-off between solution quality and computational efficiency; HC, while fast, compromises on solution accuracy; and GA, despite its adaptability, performs suboptimally in both solution quality and computational speed. These findings align with the "no free lunch" theorem, which suggests that no single metaheuristic excels universally across all problem domains. Instead, each metaheuristic offers specific advantages and trade-offs, reinforcing the notion that a combination of methods may be more effective for complex optimization tasks like CVRP, which often necessitate more than one approach.

In this context, a multi-agent system that integrates the strengths of multiple algorithms could be highly beneficial. Each metaheuristic would contribute its unique advantages at different stages of the optimization process, resulting in more robust and adaptable solutions. Looking ahead, future research will explore the dynamic integration of these four metaheuristics within a multi-agent framework. By selectively deploying and combining algorithms based on the problem's characteristics, we aim to develop more efficient and flexible approaches for tackling complex optimization challenges in logistics, transportation, and beyond.

References

[1] Ghita Bencheikh. Metaheuristics and machine learning convergence: A comprehensive survey and future prospects. *Metaheuristic and Machine Learning Optimization Strategies for Complex Systems*, pages 276–322, 2024.

[2] Marco Dorigo Vittorio Maniezzo Alberto Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Trans Syst Man Cybernetics - Part B*, 1996.

[3] Marco Dorigo and Thomas Stützle. Ant colony optimization: Overview and recent advances. In Potvin Jean-Yves Gendreau Michel, editor, *Handbook of Metaheuristics*, pages 311–351. Springer International Publishing, Cham, 2019.

[4] Dantzig G. and Ramser J. The truck dispatching problem. *Management Science, Vol. 6, No. 1*, pages 80–91, 1959.

[5] J. H. Holland. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. *MIT Press*, 1975.

[6] Rahul Kala. Chapter 15 - an introduction to evolutionary computation. In Rahul Kala, editor, *Autonomous Mobile Robots*, Emerging Methodologies and Applications in Modelling, pages 715–759. Academic Press, 2024.

[7] Luka Matijević, Vladimir Ilin, Tatjana Davidović, Tatjana Jakšić-Krüger, and Panos M. Pardalos. General vns for asymmetric vehicle routing problem with time and capacity constraints. *Computers & Operations Research*, 167:106630, 2024.

[8] Riccardo Poli, James Kennedy, and Tim Blackwell. Particle swarm optimization: An overview. *Swarm Intelligence*, 1, 10 2007.

[9] Miller C.; Tucker A.; Zemlin R. Integer programming formulations and travelling salesman problems. *Assoc. Comput. Mach.*, 1960.

[10] Jianmai Shi, Huiting Mao, Zhongbao Zhou, and Long Zheng. Adaptive large neighborhood search algorithm for the unmanned aerial vehicle routing problem with recharging. *Applied Soft Computing*, 147:110831, 2023.

[11] Yong Shi, Wenheng Liu, and Yanjie Zhou. An adaptive large neighborhood search based approach for the vehicle routing problem with zone-based pricing. *Engineering Applications of Artificial Intelligence*, 124:106506, 2023.

[12] Marduch Tadaros, Angelo Sifaleras, and Athanasios Migdalas. A variable neighborhood search approach for solving a real-world hierarchical multi-echelon vehicle routing problem involving hct vehicles. *Computers & Operations Research*, 165:106594, 2024.

[13] Krishna Veer Tiwari and Satyendra Kumar Sharma. An optimization model for vehicle routing problem in last-mile delivery. *Expert Systems with Applications*, 222:119789, 2023.

[14] Ning Xue, Ruibin Bai, Rong Qu, and Uwe Aickelin. A hybrid pricing and cutting approach for the multi-shift full truckload vehicle routing problem. *European Journal of Operational Research*, 292(2):500–514, 2021.

[15] Öncan T.; Altınel I.; Laporte G. A comparative analysis of several asymmetric traveling salesman problem formulations. *Comput. Oper. Res.*, 2009.

Data Management through the entire vaccine lifecycle: issues and challenges

Sanae ALLAM[1], Adiba EL BOUZEKRI EL IDRISSI[1] , Aicha SEKHARI[2] , Yacine OUZROUT[2]

[1] Chouaïb Doukkali University, Natl Schl Appl Sci, Sci Engineer Lab Energy, El Jadida, Morocco;

[2] DISP laboratory, University of Lumière Lyon 2, Lyon, France;

Allam.s@ucd.ac.ma

Abstract: Effective data management is essential throughout the vaccine lifecycle, encompassing research and development, manufacturing, distribution, and post-market surveillance. This paper critically examines the key challenges in managing vaccine-related data across various stages, including issues of data integration, interoperability, privacy, security, and regulatory compliance. The fragmented nature of data systems across stakeholders such as pharmaceutical companies, regulatory bodies, healthcare providers, and supply chain entities creates inefficiencies that hinder timely decision-making and vaccine delivery. Ensuring data integrity, particularly during clinical trials and regulatory submissions, is vital to maintaining public trust and avoiding approval delays. Technological advancements present transformative solutions, enhancing data integration, security, and real-time monitoring capabilities. However, disparities in infrastructure and technical expertise, particularly in low- and middle-income countries, underscore the need for global efforts to standardize and adopt these technologies. This paper highlights the importance of creating a resilient system capable of supporting the efficient, transparent, and equitable distribution of vaccines worldwide.

Keywords *: Vaccine Lifecycle, Data Management, Interoperability, Advanced technologies*

1 Introduction

The management of data throughout the vaccine lifecycle has become a critical challenge in the modern world of public health. Vaccines, which are among the most successful tools for preventing diseases and saving millions of lives globally, depend heavily on effective data management across stages ranging from research and development to post-market surveillance. The complexity of the vaccine lifecycle—spanning discovery, clinical trials, regulatory approval, manufacturing, distribution, administration, and long-term monitoring requires robust systems that ensure accuracy, transparency, and security. Recent experiences, particularly during the COVID-19 pandemic, have underscored the critical importance of efficient data management in accelerating vaccine deployment while maintaining high safety and efficacy standards [1].

At the core of these challenges lies the integration of data from multiple stakeholders, including pharmaceutical companies, regulatory bodies, healthcare providers, and governments, at different stages of the vaccine lifecycle. Each phase generates large volumes of data, from early research findings and clinical trial outcomes to manufacturing processes and distribution logistics. Additionally, the increasing reliance on digital technologies in vaccine management has introduced opportunities and challenges, particularly in maintaining data integrity and security. For instance, blockchain technology has emerged as a promising solution to transparency and traceability issues in the vaccine supply chain. Blockchain's decentralized and immutable nature offers a safeguard against data tampering, which has been a recurring problem in global vaccine logistics [2].

Ensuring data integrity, defined as maintaining and assuring the accuracy and consistency of data, is one of the most critical challenges in vaccine lifecycle management [3]. Breaches in data integrity can result in failed trials, delayed approvals, and diminished public trust. The implementation of advanced technologies, such as Laboratory Information Management Systems (LIMS) and Enterprise Resource Planning (ERP) systems, has become increasingly necessary for maintaining data integrity during manufacturing. However, integrating these systems with regulatory frameworks and achieving compliance remains a significant hurdle for many organizations.

The global nature of vaccine distribution, particularly highlighted during the COVID-19 pandemic, has introduced further complexities. For example, maintaining proper temperature conditions throughout vaccine storage and transportation is crucial for efficacy, yet inconsistencies in data monitoring across the cold chain persist. Advanced temperature monitoring technologies, including electronic sensors and chemical indicators, are integral to preserving vaccine quality during transport. However, significant gaps remain in data collection, integration, and management at various stages of the supply chain.

Effective data management also necessitates robust regulatory frameworks to ensure vaccines meet safety and efficacy standards at every stage of their lifecycle. Regulatory oversight during clinical trials, such as those seen in the Pfizer COVID-19 vaccine trials, underscores the need for accurate, complete, and well-managed data. Lapses in data integrity or monitoring during trials can delay regulatory approvals, erode public trust, and compromise patient safety.

This paper critically examines the key challenges in managing data across the vaccine lifecycle, with a focus on integration, interoperability, data integrity, and regulatory compliance. It explores technological innovations such as artificial intelligence (AI), blockchain, and IoT that are addressing these challenges while highlighting persistent gaps, particularly in low- and middle-income countries. The paper emphasizes the importance of global collaboration, investment in data infrastructure, and policy alignment to create a resilient and equitable vaccine data management system capable of responding to current and future public health challenges.

2 Methodology

To address the research question effectively, this study adopts a systematic methodology to review and analyze the existing literature. Using Scopus as the primary database, we identified over 200 papers relevant to the intersection of data management and the vaccine lifecycle. A distribution by years was created to analyze the trend of publications over time (see Figure 1). Subsequently, all the papers were classified into three categories, as shown in Table 1 (see Tableau 1).

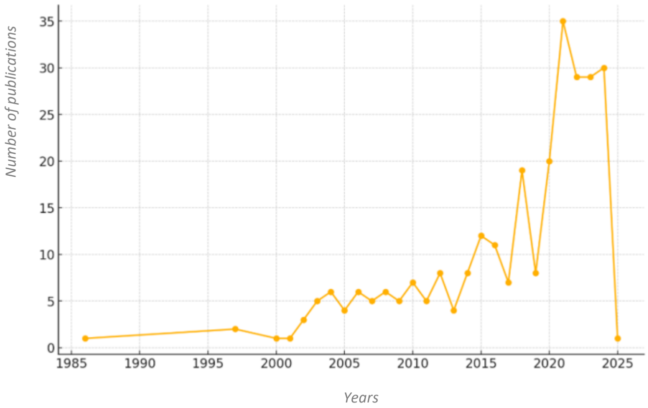


FIG. 1 – Distribution of papers by years

	Description	Number of papers
Technological Innovations	Articles focusing on AI, ML, blockchain, and other emerging technologies	266
Data Management Challenges	Articles discussing issues like GDPR compliance, data security, and interoperability	10
Case Studies	Articles providing real-world applications and case studies, including COVID-19	74

TAB. 1 – Themes Identified

These papers then underwent a rigorous selection process based on three criteria:

- Relevance to the Vaccine Lifecycle: Papers were assessed for their direct contribution to understanding or improving data management at various stages of the vaccine lifecycle, such as research, clinical trials, manufacturing, distribution, and post-market surveillance.

- Focus on Technological or Data Management Innovations: Priority was given to studies discussing the application of technologies like AI, blockchain, or data privacy frameworks in vaccine-related contexts.

- Scientific Rigor and Practical Application: Selected papers demonstrated robust methodologies and offered practical insights or real-world case studies, such as those related to COVID-19 vaccine development and distribution.

The result of this selection process was a final dataset of approximately 20 highly relevant papers, which serve as the foundation for the thematic analysis presented in this study.

3 The vaccine lifecycle : an overview

The vaccine lifecycle is a complex and multistage process that begins with the discovery of potential antigens and extends through development, clinical trials, regulatory approval, manufacturing, distribution, and post-market monitoring. Each phase generates large volumes of critical data, requiring effective management to ensure the successful development and deployment of vaccines. A robust data management system is essential at each stage to ensure safety, efficacy, and compliance with regulatory standards.

3.1 Discovery and Research

The vaccine lifecycle begins with the identification of potential pathogens or antigens capable of triggering an immune response. This phase involves extensive research, including laboratory studies, genetic sequencing, and preclinical experiments on animal models. At this stage, large amounts of biological and experimental data are generated, which must be meticulously recorded and managed. The integrity of this data is crucial for informing the next steps in the development process. Moreover, collaboration between research institutions globally is often necessary, making the integration of data across different systems and formats a significant challenge [4].

3.2 Preclinical and Clinical Trials

Once a vaccine candidate has been identified, it moves into preclinical testing, typically conducted in laboratory settings and on animals to assess safety and immunogenicity. If the results are promising, the vaccine advances to clinical trials, which are conducted in three phases on human volunteers.

- Phase I trials test the vaccine on a small group of individuals to assess its safety and immune response.
- Phase II trials involve a larger group to evaluate the vaccine's efficacy and optimal dosing.
- Phase III trials are large-scale trials that assess the vaccine's efficacy and safety in a much larger population.

Managing data in clinical trials is particularly challenging due to the sheer volume and diversity of information collected, ranging from patient demographics and health conditions to adverse event reporting and immune responses. Additionally, clinical trials are often conducted across multiple countries and institutions, necessitating robust data integration and harmonization across different systems [5] [6].This stage is also subject to stringent regulatory requirements for data collection, reporting, and transparency, as seen in issues like those raised during the Pfizer COVID-19 vaccine trials, where lapses in data integrity were exposed.

3.3 Regulatory Approval

Once a vaccine has successfully passed through clinical trials, the next critical stage is regulatory approval. Regulatory agencies, such as the U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA), and the World Health Organization (WHO), require detailed reports on the safety, efficacy, and quality of the vaccine, based on the data collected during clinical trials.

In this phase, data integrity and transparency are essential, as any discrepancies or errors in the data can delay approval and potentially undermine public trust. Regulatory submissions also require a clear audit trail, showing the origin and accuracy of all data submitted. Additionally, different regions may have varying regulatory requirements, which can further complicate the data submission process. In some cases, accelerated or emergency use authorization pathways, such as those used during the COVID-19 pandemic, allow for the parallel submission of data while trials are ongoing, adding additional layers of complexity to data management.

3.4 Manufacturing and Scale-Up

Once regulatory approval is granted, vaccines move into the manufacturing phase. This phase involves scaling up production from small-scale batches used in clinical trials to mass production for widespread distribution. Throughout this process, maintaining data integrity is critical to ensure that the vaccine produced at scale meets the same quality and safety standards as those used in clinical trials.

Manufacturers are required to document all aspects of the production process, including batch records, quality control data, and adherence to Good Manufacturing Practices (GMP). Data related to the temperature and handling conditions throughout the manufacturing process are also critical, especially for vaccines that require cold storage to remain effective. The increased complexity of supply chains during this stage further emphasizes the need for robust data management systems to track and monitor the vaccine's quality from production to distribution.

3.5 Distribution and Cold Chain Management

Once manufactured, vaccines must be distributed to various regions and healthcare facilities. Vaccines are temperature-sensitive biological products, and their potency can be compromised if they are exposed to improper storage conditions. Ensuring that vaccines remain effective during transport requires the use of cold chain systems, which maintain specific temperature ranges from the point of manufacture to the point of administration.

Cold chain data management involves monitoring the temperature of vaccines at every stage of the supply chain, from storage warehouses to transportation vehicles, and eventually to healthcare providers. Temperature excursions, or deviations from the required temperature range, must be documented and acted upon immediately to prevent the distribution of compromised vaccines. Electronic temperature monitoring devices, such as data loggers, are now widely used to track and record temperature data in real-time throughout the cold chain [7]. However, challenges remain, particularly in low- and middle-income countries (LMICs), where cold chain infrastructure may be underdeveloped [4].

3.6 Post-Market Surveillance and administration

Even after a vaccine has been distributed and administered, the lifecycle is far from over. Post-market surveillance involves monitoring the long-term safety and efficacy of the vaccine in the general population. This stage requires the collection and analysis of real-world data on adverse events, vaccine failures, and overall population health outcomes [2]. Administration systems rely heavily on the timely and accurate reporting of adverse events by healthcare providers and patients. Data integration at this stage is essential, as reports must be consolidated and analyzed across different regions and healthcare systems to detect patterns that may indicate safety concerns. Moreover, real-world data on vaccine performance is crucial for informing future vaccine development and improving existing vaccines [2]. However, data collection in this phase often suffers from underreporting and inconsistencies, posing challenges for comprehensive vaccine monitoring.

4 Key Challenges and Proposed Solutions in Vaccine Lifecycle Data Management : Literature analysis

Managing data effectively throughout the vaccine lifecycle is crucial to ensuring the successful development, production, and distribution of vaccines. However, significant challenges arise due to the complexity of integrating data across global supply chains, navigating regulatory frameworks, and ensuring compatibility with diverse healthcare systems. One of the most significant challenges is data integration. Stakeholders such as research institutions, pharmaceutical companies, regulatory bodies, and healthcare providers use different systems and formats for storing and managing data. This results in data silos that hinder the seamless exchange of information and create inefficiencies. During the COVID-19 vaccine rollout, integrating clinical trial data across multiple countries posed substantial challenges. Despite the availability of interoperability standards like HL7 and FHIR, their adoption has been slow, particularly in regions with underdeveloped healthcare IT infrastructure.

To address data integration challenges, adopting interoperability standards and modernizing IT infrastructure is critical. For instance, HL7 and FHIR facilitate the consistent exchange of data between systems, enabling real-time data sharing and collaboration. However, resource limitations and reliance on legacy systems in many regions impede progress. Public health bodies in developed countries have started implementing these frameworks, but significant gaps remain in developing regions, where technical and financial resources are constrained. Bridging these gaps is essential to enhance global data management capabilities.

Another critical challenge is ensuring data integrity across the vaccine lifecycle. Data integrity refers to maintaining accuracy, consistency, and reliability, which are essential for decision-making at every stage. Breaches in data integrity can lead to clinical trial errors, delays in regulatory approvals, and public health risks. The increasing reliance on digital tools and electronic records has improved the ability to monitor data integrity but has also introduced new risks such as cyber threats and human error. Blockchain technology offers a promising solution to these issues by providing a tamper-proof ledger for tracking data across the vaccine lifecycle. However, the widespread adoption of blockchain remains limited due to resource constraints and a lack of technical expertise.

To complement technological solutions, organizations must adopt Good Documentation Practices (GDP) and conduct regular audits to ensure that data remains reliable and consistent. For example, vaccine trials rely heavily on precise and accurate data recording, as any discrepancies can undermine confidence in a vaccine's efficacy or safety. Although technologies like blockchain can mitigate risks, further research and investment are needed to adapt these solutions to resource-constrained environments [8].

Data quality is another fundamental challenge in vaccine lifecycle management. Inconsistent or inaccurate data can lead to flawed conclusions, especially in clinical trials where decisions rely on large datasets from diverse sources [9]. For instance, during vaccine transportation, inaccurate cold chain monitoring data can result in spoiled vaccines, posing health risks and financial losses [10]. Automated systems, such as IoT-enabled temperature sensors, can reduce human error and improve data accuracy. Centralized data repositories with real-time validation mechanisms can further ensure data reliability. However, the adoption of such systems remains uneven, particularly in low-resource settings, highlighting the need for scalable and cost-effective solutions.

Regulatory compliance adds an additional layer of complexity to vaccine data management. Regulatory agencies like the FDA and EMA require extensive documentation for vaccine approval, which often varies across regions [11]. For example, the FDA's 21 CFR Part 11 and the EU's Annex 11 outline different requirements for managing electronic records and signatures. Ensuring compliance with these diverse regulations is resource-intensive and requires robust systems with transparent audit trails. Harmonization of global regulatory requirements could significantly reduce this burden, enabling more efficient vaccine approval processes.

In addressing these challenges, significant progress has been made through technological innovations and stricter adherence to regulatory guidelines. However, several gaps remain. Resource inequality between developed and developing regions limits access to advanced data management systems [12]. The slow adoption of interoperability standards and blockchain technology continues to impede progress. Additionally, the lack of regulatory alignment across regions creates unnecessary complexity in vaccine approvals. Future efforts should focus on developing cost-effective and scalable solutions for data management, accelerating the adoption of global interoperability standards, and exploring decentralized systems like blockchain for broader application in vaccine supply chains and regulatory compliance. Addressing these gaps will be critical to creating a robust and efficient vaccine lifecycle data management system that can respond effectively to global health challenges.

Managing data effectively throughout the vaccine lifecycle is critical to ensure the successful development, production, and distribution of vaccines [13]. However, several key challenges arise at different stages of the lifecycle, particularly as the complexity of managing data across global supply chains, regulatory frameworks, and healthcare systems increases. These challenges include data integration, ensuring data integrity, safeguarding data privacy and security, maintaining data quality, and meeting diverse regulatory compliance requirements.

5 Discussion & Future Perspectives

The vaccine lifecycle, spanning research, development, manufacturing, distribution, and post-market surveillance, is heavily dependent on efficient and accurate data management. In an era defined by rapid technological advancements and global health crises like COVID-19, managing vaccine-related data has never been more critical. This discussion explores the challenges and technological innovations that define the current landscape, proposing actionable solutions while identifying persistent gaps.

A significant challenge in the vaccine lifecycle is the fragmentation of data across multiple stakeholders, including research institutions, pharmaceutical companies, regulatory agencies, and healthcare providers. Each stakeholder employs diverse systems and standards, leading to silos of information that impede data integration. These silos not only slow down decision-making but also increase the likelihood of errors and redundancies. The adoption of interoperability standards such as HL7 and FHIR has shown potential to bridge these gaps, enabling seamless data exchange across platforms. For instance, during the COVID-19 pandemic, these standards played a vital role in accelerating vaccine approvals by facilitating real-time data sharing among global stakeholders. However, their adoption remains uneven, particularly in low- and middle-income countries (LMICs), where outdated systems and limited resources continue to hinder progress. Addressing these disparities requires global investment in healthcare IT infrastructure and coordinated policy efforts to standardize data management practices across regions.

Ensuring data integrity throughout the vaccine lifecycle presents another critical hurdle. Data integrity defined as the accuracy, consistency, and reliability of data—is indispensable for making informed decisions at every stage, from preclinical trials to post-market surveillance. However, breaches in data integrity can lead to catastrophic outcomes, such as delays in vaccine approvals or diminished public trust in vaccination campaigns. For example, inconsistencies in clinical trial data for COVID-19 vaccines raised concerns about transparency and accountability during regulatory reviews. Blockchain technology, with its ability to provide an immutable and transparent ledger, has emerged as a promising solution. By ensuring that all transactions and data entries are securely recorded, blockchain can enhance trust and traceability in vaccine data management. Despite its potential, blockchain adoption remains limited, with implementation challenges including high costs and the need for specialized expertise. To overcome these barriers, pilot programs demonstrating the efficacy of blockchain in vaccine logistics and clinical trials should be scaled up, particularly in regions with well-developed healthcare systems.

Data quality is an equally pressing concern. The vaccine lifecycle relies on vast datasets, including clinical trial results, patient demographics, and logistical information such as cold chain monitoring. Inconsistent or inaccurate data can undermine vaccine efficacy and safety assessments, leading to flawed public health decisions. For example, temperature deviations during vaccine transport, if not properly recorded, can result in compromised vaccine potency, posing risks to public health and financial losses. Innovations in automated data collection, such as IoT-enabled sensors, are addressing these issues by providing real-time monitoring and reducing the risk of human error. However, these solutions are often inaccessible in resource-limited settings, where manual data entry remains the norm. Bridging this gap requires investments in scalable, cost-effective technologies that can be deployed in LMICs without requiring extensive technical expertise.

Regulatory compliance adds another layer of complexity to vaccine data management. Regulatory agencies such as the FDA and EMA demand comprehensive documentation to ensure vaccine safety and efficacy. However, navigating the varied requirements of different jurisdictions can be resource-intensive and time-consuming, particularly for manufacturers operating globally. Digital systems with robust audit trails can streamline compliance processes by ensuring traceability and transparency in data submissions. Additionally, greater harmonization of regulatory requirements across regions would reduce the burden on manufacturers and accelerate vaccine availability during health emergencies. The COVID-19 pandemic highlighted the urgent need for such harmonization, as inconsistent regulatory frameworks slowed the global rollout of vaccines.

Despite significant advancements, gaps remain in ensuring that technological solutions are accessible, scalable, and aligned with the realities of healthcare systems worldwide. For example, while AI and ML are transforming clinical trial analysis and pharmacovigilance, their deployment in LMICs is limited by resource constraints and lack of expertise. Similarly, blockchain and IoT technologies, while promising, require substantial initial investments that are often out of reach for resource-constrained health systems.

The future of vaccine data management lies at the intersection of technology, policy, and collaboration. While challenges in data integration, integrity, quality, and compliance remain significant, innovations such as AI, blockchain, IoT, and cloud computing offer powerful tools to address these issues. However, the success of these technologies depends on their accessibility and integration into a broader framework of organizational and policy reforms. By fostering global collaboration and investing in robust data infrastructures, the global health community can build a resilient system that ensures the safety, efficacy, and equitable distribution of vaccines. Lessons learned from the COVID-19 pandemic highlight the urgency of this endeavor, underscoring the need for proactive measures to prepare for future health crises. Through a coordinated effort, vaccines can continue to play their vital role in safeguarding public health globally.

6 Conclusion

Effective data management is the backbone of the vaccine lifecycle, encompassing research, development, manufacturing, distribution, and post-market surveillance. As this paper has demonstrated, managing the vast amounts of data generated at each stage of this lifecycle is both a critical necessity and a significant challenge. The fragmentation of data systems across various stakeholders—research institutions, pharmaceutical companies, regulatory agencies, healthcare providers, and supply chain entities—creates silos that hinder integration, interoperability, and decision-making. Addressing these challenges requires not only technological advancements but also robust frameworks for collaboration, standardization, and policy alignment.

The COVID-19 pandemic has brought the importance of efficient data management into sharp focus, exposing both the strengths and weaknesses of current systems. Rapid vaccine development and deployment during the pandemic underscored the potential of advanced technologies like artificial intelligence (AI), blockchain, the Internet of Things (IoT), and cloud computing to enhance data management processes. AI and machine learning have revolutionized data analysis in clinical trials and post-market surveillance, enabling faster identification of safety signals and more accurate predictions. Blockchain technology, with its immutable ledger, offers unparalleled transparency and traceability, addressing critical concerns around data integrity and security. IoT-enabled sensors have improved cold chain monitoring, ensuring that vaccines remain effective during storage and transportation. However, while these technologies offer transformative solutions, their implementation is uneven across regions, particularly in low- and middle-income countries (LMICs), where disparities in infrastructure, technical expertise, and resources persist.

One of the key findings of this study is that the adoption of interoperability standards like HL7 and FHIR is essential for bridging data silos and enabling seamless collaboration among stakeholders. These standards facilitate the exchange of information across diverse platforms, promoting efficiency and reducing redundancies. However, the slow and inconsistent adoption of these standards, particularly in resource-constrained settings, remains a significant barrier. Global initiatives to promote the adoption of interoperability frameworks, supported by investments in healthcare IT infrastructure, are critical to overcoming this challenge. Without such efforts, the global health community risks perpetuating the inequities that currently hinder vaccine distribution and accessibility.

Ensuring data integrity has emerged as another critical focus area. The accuracy, consistency, and reliability of data are fundamental to every stage of the vaccine lifecycle, from preclinical research to regulatory approval and post-market monitoring. Breaches in data integrity, whether due to cyber threats, human error, or outdated systems, can have catastrophic consequences, including delays in vaccine approvals, loss of public trust, and compromised health outcomes. Blockchain technology holds significant promise in addressing these challenges, providing a secure and transparent mechanism for recording data transactions. However, as highlighted in this paper, the adoption of blockchain across the vaccine lifecycle remains limited due to high implementation costs, the need for specialized expertise, and regulatory uncertainties. To unlock the full potential of blockchain, pilot programs and case studies must be scaled up, particularly in developed regions with advanced healthcare systems, to demonstrate its efficacy and cost-effectiveness.

The quality of data collected and managed throughout the vaccine lifecycle is another pressing concern. Inconsistent or inaccurate data can undermine vaccine efficacy and safety assessments, leading to flawed public health decisions and financial losses. For instance, lapses in cold chain monitoring can compromise vaccine potency, posing risks to public health. Automated systems, such as IoT-enabled temperature sensors, have shown promise in mitigating these risks by providing real-time data and reducing human error. However, the uneven adoption of such systems, particularly in LMICs, underscores the need for scalable and cost-effective solutions that can be deployed in resource-limited settings. Investments in centralized data repositories with real-time validation mechanisms can further enhance data reliability and support informed decision-making across the vaccine lifecycle.

Regulatory compliance adds another layer of complexity to vaccine data management. Regulatory agencies like the FDA, EMA, and WHO play a critical role in ensuring the safety, efficacy, and quality of vaccines. However, navigating the diverse and often conflicting requirements of these agencies is a resource-intensive process that can delay vaccine approvals and distribution. The COVID-19 pandemic highlighted the urgent need for greater harmonization of regulatory requirements across regions. Aligning regulatory frameworks can reduce inefficiencies, accelerate vaccine availability, and ensure that lifesaving interventions reach populations in need without unnecessary delays. Digital systems with robust audit trails can further streamline compliance processes, providing transparency and traceability in data submissions.

While technological innovations offer powerful tools to address the challenges of vaccine data management, this study highlights that technology alone is not sufficient. Effective data management requires a holistic approach that combines technological advancements with organizational reforms, policy alignment, and global collaboration. Governments, healthcare providers, pharmaceutical companies, and regulatory bodies must work together to develop standardized, interoperable systems that facilitate seamless data sharing and ensure compliance with global regulatory requirements. Equally important is the need for targeted investments in healthcare IT infrastructure and capacity building in LMICs. By addressing resource disparities and enhancing technical expertise, the global health community can ensure equitable access to vaccines and strengthen global health security.

This study also underscores the importance of ethical and legal considerations in vaccine data management. As technologies like AI and blockchain become more prevalent, questions around data privacy, consent, and ownership must be addressed. Regulations like the General Data Protection Regulation (GDPR) provide a foundational framework for protecting sensitive health information, but their implementation often lags behind technological developments. Strengthening enforcement mechanisms and promoting transparency in data collection and usage are critical to maintaining public trust and ensuring the ethical management of vaccine-related data.

Looking ahead, the future of vaccine data management lies at the intersection of technology, policy, and global cooperation. The lessons learned from the COVID-19 pandemic highlight the urgency of building resilient data management systems that can respond swiftly to emerging public health crises. By leveraging the latest technological advancements, aligning regulatory frameworks, and fostering collaboration across sectors, the global health community can create a robust and equitable system for managing vaccine data. Such a system will not only enhance the efficiency and transparency of vaccine development and distribution but also ensure that vaccines continue to play a vital role in safeguarding public health worldwide.

In conclusion, addressing the challenges of vaccine data management is a shared responsibility that requires collective action and sustained investment. By prioritizing data integrity, quality, and interoperability, and by embracing innovative solutions, the global health community can overcome existing barriers and build a future where vaccines are accessible, effective, and trusted by all.

7 Referencing

[1] L. H. M. Van De Burgwal, C. D. S. Ribeiro, M. B. Van Der Waal, and E. Claassen, “Towards improved process efficiency in vaccine innovation: The Vaccine Innovation Cycle as a validated, conceptual stage-gate model,” *Vaccine*, vol. 36, no. 49, pp. 7496–7508, Nov. 2018, doi: 10.1016/j.vaccine.2018.10.061.

[2] C. Subbarao, B. Saidamma, P. V. Narasimha Raju, C. Srinivas, I. K. Varghese, and B. Shiva, “Role of Block Chain Technology in Improving Security and Efficiency in Pharmaceutical Supply Chains,” in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, Coimbatore, India: IEEE, Aug. 2024, pp. 6–12. doi: 10.1109/ICoICI62503.2024.10696562.

[3] P. D. Thacker, “Covid-19: Researcher blows the whistle on data integrity issues in Pfizer’s vaccine trial: Video 1,” *BMJ*, p. n2635, Nov. 2021, doi: 10.1136/bmj.n2635.

[4] P. D. Thacker, “Covid-19: Researcher blows the whistle on data integrity issues in Pfizer’s vaccine trial: Video 1,” *BMJ*, p. n2635, Nov. 2021, doi: 10.1136/bmj.n2635.

[5] L. H. M. Van De Burgwal, C. D. S. Ribeiro, M. B. Van Der Waal, and E. Claassen, “Towards improved process efficiency in vaccine innovation: The Vaccine Innovation Cycle as a validated, conceptual stage-gate model,” *Vaccine*, vol. 36, no. 49, pp. 7496–7508, Nov. 2018, doi: 10.1016/j.vaccine.2018.10.061.

[6] I. Ilin, A. Levina, and K. Frolov, “Innovative Ecosystem Model of Vaccine Lifecycle Management,” *J. Open Innov. Technol. Mark. Complex.*, vol. 8, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/joitmc8010005.

[7] R. Anderson *et al.*, “Supporting immunization programs with improved vaccine cold chain information systems,” in *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, San Jose, CA: IEEE, Oct. 2014, pp. 215–222. doi: 10.1109/GHTC.2014.6970284.

[8] M. Durá *et al.*, “Blockchain for Data Originality in Pharma Manufacturing,” *J. Pharm. Innov.*, vol. 18, no. 4, pp. 1745–1763, Dec. 2023, doi: 10.1007/s12247-023-09748-z.

[9] O. Choudhury, I. Sylla, N. Fairoza, and A. Das, “A Blockchain Framework for Ensuring Data Quality in Multi-Organizational Clinical Trials,” in *2019 IEEE International Conference on Healthcare Informatics (ICHI)*, Xi’an, China: IEEE, Jun. 2019, pp. 1–9. doi: 10.1109/ICHI.2019.8904634.

[10] K. Bollaerts *et al.*, “The role of real-world evidence for regulatory and public health decision-making for Accelerated Vaccine Deployment- a meeting report,” *Biologicals*, vol. 85, p. 101750, Feb. 2024, doi: 10.1016/j.biologicals.2024.101750.

[11] L. Scheppler, N. De Clercq, M. McGoldrick, and J. Dias, “Regulatory Harmonization and Streamlining of Clinical Trial Applications globally should lead to faster clinical development and earlier access to life-saving vaccines,” *Vaccine*, vol. 39, no. 5, pp. 790–796, Jan. 2021, doi: 10.1016/j.vaccine.2020.11.077.

[12] A. Guignard, N. Praet, V. Jusot, M. Bakker, and L. Baril, “Introducing new vaccines in low- and middle-income countries: challenges and approaches,” *Expert Rev. Vaccines*, vol. 18, no. 2, pp. 119–131, Feb. 2019, doi: 10.1080/14760584.2019.1574224.

[13] S. Plotkin, J. M. Robinson, G. Cunningham, R. Iqbal, and S. Larsen, “The complexity and cost of vaccine manufacturing – An overview,” *Vaccine*, vol. 35, no. 33, pp. 4064–4071, Jul. 2017, doi: 10.1016/j.vaccine.2017.06.003.

Improvement of Accuracy and Flexibility in 3D Cardiac Models through AI and Operational Research

Aya Abiya[1], Fatima Bouyahia[1], Mouhcine El Mardouli[2], Drissi Boumzebra[2], Ahmed Aamouche [2]

[1] Systems Engineering & Applications Laboratory LISA, Cadi Ayyad UniversityMarrakesh
ayaabiya11@gmail.com,f.bouyahia@uca.com,a.aamouche@uca.com
[2] Health & Bioscience Laboratory, Cadi Ayyad University-Marrakesh
mouhcinemard@gmail.com,dr.boumzebra@uca.ma

***Keywords**: Optimization, Operational Research, Cardiovascular, Tetralogy of Fallot, Artificial Intelligence, 3D Printing, Modelling, Simulation, Decision Making.*

1 Introduction

In this study, we explore the use of advanced modeling and simulation tools to enhance clinical decision-making and the training of future surgeons in the field of cardiology. Tetralogy of Fallot, a complex congenital heart defect, presents significant challenges due to the anatomical variability of patients and the complexity of the required surgical interventions. To address these challenges, this research proposes an innovative approach to 3D modeling, leveraging Artificial Intelligence (AI) and Operations Research (OR).

Traditionally, the diagnosis and treatment of heart diseases rely on conventional imaging techniques, which are often inadequate for the complex interventions required for Structural Heart Diseases (SHD). The integration of imaging into periprocedural guidance has marked significant progress, yet the visualization and manipulation of precise and flexible models remain crucial. In response to this need, our approach combines two industrial processes optimized through OR to create realistic and malleable heart models, tailored to the specific conditions of the patients.

AI played a central role in integrating radiological data, notably MRI images, with patient-specific information such as medical history and genetic factors. This integration enabled the generation of highly accurate and malleable 3D models of hearts affected by Tetralogy of Fallot. These models not only enhance the understanding of cardiac diseases but also provide essential training support for surgeons before real surgical interventions. By practicing on faithful replicas of patients' hearts, surgeons can improve their preparation and the safety of operations.

To validate the effectiveness of our approach, we conducted a simulation during a medical conference, bringing together 150 doctors from 60 different nationalities. The results were very positive, highlighting the realism and utility of the models for training and surgical practice.

This study demonstrates how the integration of AI and OR can transform decision-making in cardiovascular medicine by providing accurate, malleable, and usable heart models for simulations. These models are manufactured in an optimized manner in terms of cost, time, and quality, thus addressing the challenges encountered and opening new perspectives for the treatment of complex cardiac diseases .

2 Bibliography citation

2.1 Heart Modeling and Simulation

The 2020 study by Hon, Hussein, Osamihonjo, and Yoo revealed decreased interest in cardiac surgery among medical students. However, the use of 3D-printed heart models for simulation significantly impacted their career choices, potentially increasing interest in cardiac surgery. [1]

In the research conducted by Anwar & al., (2017), the authors used three-dimensional (3D) printing technology in patients diagnosed with congenital heart disease. Their goal was to represent complex anatomical structures accurately, plan surgical interventions, and educate both medical trainees and patients. The cases examined included individuals of different age groups, from infants to adults, all with congenital heart conditions. Various pathologies were explored, covering intracardiac malformations, vascular anomalies, and airway irregularities. [2]

The 2013 study by D.B.S. Tam & al., illustrates the use of 3D printing to create a replica of an aortic aneurysm. This modelhelped guide decision-making and device selection for endovascular repair of the aneurysm. 3D printing allowed a detailed analysis of the patient’s anatomy, facilitating the selection of appropriate devices. This improved preoperative planning and reduced potential complications. Surgeons can use the 3D model to simulate different treatment scenarios and evaluate their effectiveness. This allows for informed decision-making, minimising risks, and optimising patient outcomes. In sum, 3D printing offers precise and personalised visualisation of the patient’s vascular anatomy, facilitating decision-making. [3]

The 2014 study by Schmauss & al., evaluates the use of 3D-printing in cardiac surgery and interventional cardiology based on the experience of a single medical center from 2006 to 2013. It demonstrates the use of 3D models for perioperative planning and simulation of cardiovascular procedures. Eight representative cases were selected to show the usefulness of 3D models in accurately visualising cardiovascular anatomy. The study concludes that 3D printing is a feasible and valuable approach to improve planning and simulation in pediatric and adult cardiac surgery and interventional cardiology, paving the way for future studies to quantify its potential benefits better. [4]

The study by Farooqi & al. analysed 3D printing methods, including FDM (Fused Deposition Modeling) and stereolithography. These techniques transform a 3D virtual object into a physical object. FDM printers deposit layers of liquefied thermoplastic material, while stereolithographic printers use a laser to solidify layers of liquid photopolymer. [5]

3D printers, such as FDM, PolyJet, and stereolithographic models, offer a range of capabilities, including multicolour and multi-material printing. FDM printers are more economically accessible, while PolyJet and stereolithographic printers offer finer layer resolution. The choice of the ideal printer depends on cost and specific purpose. Using 3D-printed models with Agilus materials was effective for surgical planning in children with Raghib syndrome. A 2020 study designed a model of complex congenital heart disease to facilitate the simulation of a surgical procedure. Once the 3D virtual file is prepared, it is printed layer by layer. Overhanging parts are generally printed with a support material to maintain their position. [6]

The 2016 study by Crafts & al., highlights that the initial costs of adopting 3D printing are often considered prohibitive, which could hinder its widespread adoption in otolaryngology. However, prices continue to fall, and it has been demonstrated that using 3D-printed materials can result in cost savings. Regarding medical applications, the limited number of materials approved by the Food and Drug Administration results in higher material costs. Although the materials used for 3D-printing of educational models are becoming increasingly accessible, many educators are consistently concerned that there is no proper substitute for human tissues. Nevertheless, the use of 3D-printed models could potentially reduce the need to acquire cadaver bones for training.[7]

The research by Yiting & al. (2019) examines the application of 3Dprinting in structural heart diseases. It highlights how this innovative technology improves understanding of anatomy, aids in preoperative planning, and facilitates surgical simulation. The study highlights the advantages of 3D printing for customising treatments and reducing surgical risks while discussing future progress and challenges. Despite the potential limitations of 3D printing, such as the quality and precision of printed anatomical models, continuous technological improvement and increasing practitioner expertise offer ways to mitigate these limitations. [8]

In 2018, Lau &Sun studied 3D printing in pediatric and congenital cardiac surgery. Despite the advantages of this technology, such as the creation of precise anatomical models, they also highlight the challenges associated with its implementation, such as the complexity of modelling cardiac structures in pediatric patients, technological constraints, and substantial costs. Thus, despite the potential of 3D printing, its limitations must be carefully considered for optimal use. [9]

2.2 Use of Operations Research in Medicine

In 2017, Muge Capan and colleagues examined the application of Operational Research (OR) in healthcare, focusing on how mathematical methods can address complex challenges within the medical field. The goal is to enhance medical decision-making and healthcare organization by providing data-driven solutions while fostering interdisciplinary collaboration. The study highlights the use of OR methods, such as linear programming, Markov decision processes, and advanced modeling techniques, often combined with statistical tools, to optimize healthcare practices.

The research identifies several key areas where OR has been successfully applied, including:

- Medical Treatment Decision Support

OR aids in optimizing treatment choices and diagnostic testing by considering factors such as costs, clinical outcomes, and patient preferences. The study emphasizes how OR can assist in making informed decisions when clinical evidence is insufficient, offering mathematical models that assess different treatment options based on risk and cost considerations.

Example : In prostate cancer treatment following prostatectomy, researchers used Markov modeling to simulate cancer progression and determine the optimal treatment strategy. The results demonstrated that incorporating personalized genomic tests could significantly improve patient outcomes by tailoring treatments to individual risks.

- Optimizing the Timing of Diagnosis

OR techniques can also optimize the timing of diagnostic tests and interventions based on patient risk and associated costs.

Example : In breast cancer screening, OR methods were applied to determine when a patient should move from mammography to biopsy, balancing the risks of false positives and delayed diagnoses, while considering factors such as the patient’s age and financial constraints.

- Optimizing the Timing of Therapeutic Interventions

OR supports decisions regarding the optimal time to initiate treatments, particularly in cases with long-term risks and uncertainties.

Example : For patients with type 2 diabetes, OR was used to model the optimal timing for initiating statin therapy, aiming to maximize quality-adjusted life years (QALYs) while accounting for cholesterol levels and cardiovascular risk.

- Organ Transplantation Optimization

The study also explores the role of OR in organ transplantation, focusing on optimizing organ distribution and donor-recipient matching to ensure a fairer and more efficient allocation of limited resources. [10]

In the 2004 study by Burke et al. and the 2012 study by Lim et al., the authors examine the application of operations research techniques in workforce scheduling, with a particular focus on nurse rostering. This area has garnered significant attention due to the complexity of designing schedules that can adapt to fluctuations in demand for healthcare services. The scheduling problem involves various constraints, including the need to distinguish between permanent and temporary staff, ensure adequate holiday coverage, balance day, night, and weekend shifts, and accommodate individual preferences. To address these challenges, several methods, including linear programming, mixed-integer programming, goal programming, and constraint programming, have been developed for generating optimal nurse rosters (Melo, 2012). [11]

In their 2008 study, Green and Savin explored the issue of delays in medical appointments, a common challenge in healthcare settings. By applying queueing theory, they aimed to model and reduce these delays, ultimately improving the efficiency of the appointment system and reducing patient waiting times. They identify several factors contributing to delays, such as variability in consultation durations, unpredictable patient arrivals, and poor scheduling practices. To address these issues, they propose several strategies, including optimizing appointment spacing based on estimated consultation times, introducing flexible appointment slots, prioritizing urgent cases, and using simulation models to dynamically adjust scheduling. Their findings demonstrate that these approaches can significantly reduce delays and improve resource utilization, highlighting the importance of queueing theory in enhancing the effectiveness of medical appointment systems. [12]

In 2008, Gupta and Denton examined the challenges and opportunities related to appointment scheduling in healthcare. They emphasize the complexities involved in managing medical appointments, such as the variability of consultation durations, unpredictable patient arrivals, and the limited availability of healthcare professionals. The authors explore various techniques, including queueing theory, optimization methods, and simulation models, to address these issues and improve scheduling efficiency. They also identify challenges like no-shows, last-minute cancellations, and the management of emergency appointments. Despite these hurdles, the study suggests that advances in dynamic scheduling, information technology, and predictive modeling offer valuable solutions for enhancing the appointment system in healthcare. Therefore, while the scheduling process remains complex, the authors argue that careful application of these methods can lead to significant improvements in efficiency and patient care. [13]

This article emphasizes the critical role of artificial intelligence (AI) and operations research (OR) in cardiology decision-making. AI was used to create detailed 3D models of hearts affected by Tetralogy of Fallot, a complex heart disease. These models, evaluated by 125 doctors from 60 nationalities, were praised for their realism and utility in surgical practice. They also enhance medical students' understanding due to their representation of blood circulation and internal details while utilizing the DMADV approach, which stands for Define, Measure, Analyze, Design, and Verify. This methodology is a part of the Six Sigma framework and is used for developing new products or processes or significantly enhancing existing ones. Additionally, operations research methods were applied to optimize the production processes of these models, ensuring cost-efficiency, quality, and timely production. By integrating OR techniques such as linear programming and simulation, the production of the heart models was streamlined, addressing key operational challenges and improving overall effectiveness. This combination of AI and OR demonstrates a transformative approach in cardiovascular medicine, enhancing both clinical decision-making and surgical training.

3 Methodology

3.1 Problem identified

Based on the information and articles reviewed, the problem statement can be framed as follows: "How can Artificial Intelligence and 3D printing technologies be utilized to enhance the understanding and treatment of complex heart conditions, such as Tetralogy of Fallot, while addressing challenges like the rigidity of models, high costs, the ongoing need for technological advancement, and the necessity for increased practitioner expertise ?"

3.2 Measuring the Impact

The articles reviewed in the Literature Survey section highlight several key points regarding the use of 3D heart models in medical training and surgery:

- 3D heart models reduce the need for cadaver hearts in training.
- They enhance student comprehension and facilitate learning.
- 3D models are valuable tools for perioperative planning and simulating cardiovascular procedures.

Additionally, the use of 3D-printed heart models has been shown to significantly influence medical students' career decisions, potentially fostering greater interest in cardiac surgery. These models offer a precise, personalized visualization of a patient's vascular anatomy, which aids in decision-making and helps reduce surgical risks.

However, despite the advantages, 3D printing does have limitations that must be addressed for optimal use. One notable challenge is the rigidity of the models, which can impede the accurate replication of complex cardiac structures, particularly in pediatric patients. To overcome these limitations and fully realize the potential of 3D printing in cardiac surgery, ongoing technological advancements and the development of greater practitioner expertise are essential.

According to the literature review, most 3D printing process outputs tend to be rigid. However, the industry offers a variety of methods, each with advantages and disadvantages. Injection moulding is one such process that could be relevant in that case.

Criteria	3D printing	Molding
Accuracy of internal details	High precision for internal cavities and complex structures (such as malformations)	Limited: difficult to reproduce very fine internal details without multiple molds
Flexibility	Very flexible: easily modifiable and adjustable	Less flexible: would require a new mold for modifications
Realistic texture and feel	Less realistic than casting materials, but possible with post-processing	Very realistic, especially with materials like silicone that mimic human tissue
Surface finish	May leave visible layers, requiring post-processing to smooth	Excellent finish, very smooth and realistic for surgical applications
Production speed	Fast, especially for prototypes or small series	Slow: Mold making and production can take time
Sustainability	Good mechanical properties but less flexible than silicone or other realistic materials	Materials like silicone or resins can be more durable and realistic
Mass production	Less suitable for large-scale mass production	Ideal for mass production once the mold has been created
Personalization	Very high customization, with easy modifications	Less flexible: changes require the creation of new molds

Cost per series model	Higher: Each impression	Much lower cost per model after mold creation

TAB. 1 – 3D printing vsInjection moulding

Both techniques boast commendable precision, yet they entail intricate technology and extensive manufacturing durations, particularly 3D printing, given its reliance on industrial processes. While 3D modelling of the heart holds promise for educational purposes and simulating cardiac surgeries, the demand from faculties and hospitals surpasses what 3D printing can efficiently provide in terms of time and pace despite printer redundancy, which escalates costs. On the other hand, opting for moulding could accelerate production rates and decrease turnaround times, albeit at the expense of requiring a steel mould, thereby adding to the overall costs.

3.3 Design

Artificial Intelligence (AI) played a significant role in this work, particularly in the field of cardiology. AI was utilised to integrate two types of data: radiological data, including images and scans of the heart, and patientspecific information, such as medical history and genetic factors. This integration enabled the creation of highly accurate 3D models of hearts affected by Tetralogy of Fallot, a complex congenital heart defect. (see Figure 1)

The use of AI aims also to enhance the rendering of these 3D models. This involved refining the visual representation of the models to make them as realistic and detailed as possible. The goal was to create a rendering that was not only accurate but also flexible and malleable, much like the actual human heart .

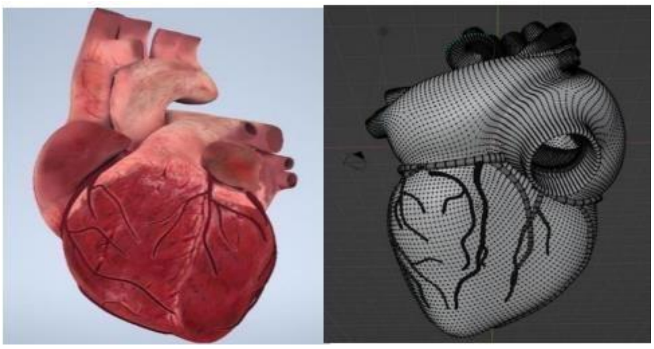


FIG. 1 – 3D modeling of the heart with a TOF.

To achieve this, AI was combined with another industrial process. This combination allowed to overcome the limitations and drawbacks inherent in these two methods when used separately. The result was the successful creation of detailed and flexible heart models (seen Fig. 2 and Fig 3).

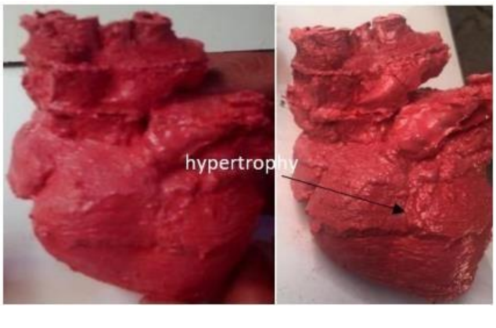


FIG. 2 – Mould and sub-moulds to facilitate demoulding(**) Internal shell before smoothing, A: shell of a healthy heart, B: diseased heart with shrinkage (Stenosis), C: (VSD).

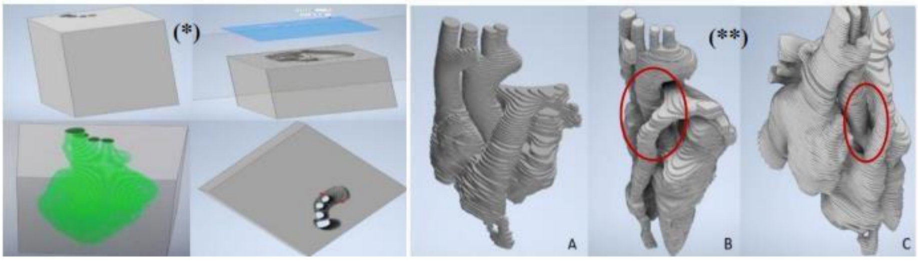


FIG. 3 – Under external mould before and after smoothing, C: Holy shell and D: Shell of the diseased heart.

These 3D heart models have proven extremely useful in the medical field. They have been used to assist doctors in making informed decisions about surgical interventions for patients with Tetralogy of Fallot. By providing a detailed and accurate visual representation of the affected heart, these models allow doctors to understand the specificities of each case better and plan the surgical interventions more effectively. (see Figure 4)

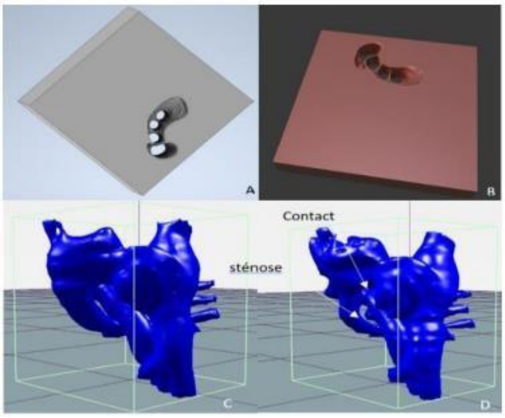


FIG. 4 – Heart in 3d with tetralogy of Fallot (Result of the method).

3.4 Implementation of Operations Research Methods and Validation of Results

In this study, we applied operations research (OR) methods to optimize the production and efficiency of 3D heart models used for pre-surgical simulations. To assess the effectiveness of these models, we focused on two key aspects: optimizing production time and ensuring model quality.

Before the medical conference, we conducted simulations to streamline the production process, with a particular focus on reducing production time. During the conference, we confirmed the results of this time optimization. The initial model took 10 hours and 45 minutes to produce, while subsequent models were created in just 1 hour and 45 minutes, confirming the success of the time optimization.

In addition to time efficiency, the conference also provided an opportunity to validate the quality of the models. During the simulation, 150 doctors evaluated whether the models accurately represented the disease, particularly Tetralogy of Fallot, and whether the tissue texture closely resembled real human tissue. The feedback from the doctors confirmed that the models were highly effective for training purposes, enabling them to better visualize complex cardiovascular structures and prepare for actual surgeries.(See Figure 5)

This dual validation—focused on both time optimization and model quality—demonstrates the practical potential of 3D heart models in enhancing surgical planning and medical training.

3.4.1 Optimization of Production Processes via Linear Programming

We used linear programming to minimize the production costs of the 3D heart models while adhering to time and quality constraints. The objective function was to minimize the total costs (C), which include material costs (Cm) and labor costs (Cl). Constraints included production time (Tp) and quality (Q).

$$\text{Minimize } C = C_m + C_l \qquad (1)$$

3.4.2 Simulation and Result

By integrating these operations research methods and validating the results through rigorous simulations and feedback, this study demonstrates substantial improvements in the production and application of 3D heart models in cardiology. These advancements not only optimize the manufacturing process but also ensure the practical effectiveness of the models for medical training and surgical preparation before undertaking real interventions. Thanks to the combination of molding and 3D printing, we significantly reduced production time after the first model. The initial heart model was created in 10 hours and 45 minutes, while subsequent models were produced in just 1 hour and 45 minutes, illustrating how process optimization can make a significant difference in production.

Result

The doctors used the models to simulate surgical interventions in a controlled environment before performing actual surgeries. This simulation allowed testing the flexibility and realism of the models, providing valuable data on their performance.

Results:

- 82% of the participants stated that the internal details of the model resembled a real heart.
- 88% noted the accuracy of the external characteristics.
- 91% believed that the model could facilitate surgical practice.
- 75% found that the material resembled natural heart tissue.
- 100% agreed that the model improved medical students' understanding.

Although conducted in a non-clinical setting, this simulation provided essential insights into how these models can be effectively used in medical training and surgical preparation. It also highlighted the advantages of malleable heart models for practical learning and clinical decision-making.

FIG. 5 – 3D heart test of more than 125 doctors from 60 nationalities.

4 Conclusion

This study highlights the significant advancements achieved through the integration of operations research (OR) methods and artificial intelligence (AI) in the production and application of 3D heart models. By optimizing production processes and validating the results through rigorous simulations and professional feedback, we have demonstrated substantial improvements in both the manufacturing efficiency and practical effectiveness of these models.

The simulations conducted, including the time optimization prior to the medical conference and the quality validation during the conference, provided critical insights into the accuracy, realism, and utility of the 3D heart models. The overwhelmingly positive feedback from 150 doctors of 60 nationalities underscores the models' potential in enhancing surgical training and preparation.

Furthermore, the combination of molding and 3D printing technologies has proven to significantly reduce production times while maintaining high-quality standards. This process optimization enabled the creation of initial heart models in 10 hours and 45 minutes, with subsequent models produced in just 1 hour and 45 minutes, illustrating the impactful role of OR in streamlining complex production processes.

Overall, the integration of OR and AI has not only optimized the production of 3D heart models but also ensured their practical utility in medical training and surgical planning. These advancements open new perspectives for the treatment of complex cardiac diseases, fostering better clinical outcomes and enhancing the educational experience for medical professionals.

5 References

[1] Hon, N.W.L., Hussein, N., Honjo, O., Yoo, S.-J., Evaluating the Impact of Medical Student Inclusion Into Hands-On Surgical Simulation in Congenital Heart Surgery. Journal of Surgical Education 78, 207–213. (2021). https://doi.org/10.1016/j.jsurg.2020.06.023

[2] Anwar, S., Singh, G.K., Varughese, J., Nguyen, H., Billadello, J.J., Sheybani, E.F., Woodard, P.K., Manning, P., Eghtesady, P., 3D Printing in Complex Congenital Heart Disease. JACC: Cardiovascular Imaging 10, 953–956. (2017). https://doi.org/10.1016/j.jcmg.2016.03.013

[3] Tam, M.D.B.S., Laycock, S.D., Brown, J.R.I., Jakeways, M., 3D Printing of an Aortic Aneurysm to Facilitate Decision Making and Device Selection for Endovascular Aneurysm Repair in Complex Neck Anatomy. Journal of Endovascular Therapy 20, 863–867. (2013). https://doi.org/10.1583/13-4450MR.1

[4] Schmauss, D., Haeberle, S., Hagl, C., Sodian, R., Three-dimensional printing in cardiac surgery and interventional cardiology: a single-centre experience. European Journal of Cardio-Thoracic Surgery 47, 1044–1052. (2015). https://doi.org/10.1093/ejcts/ezu310

[5] Farooqi, K.M., Saeed, O., Zaidi, A., Sanz, J., Nielsen, J.C., Hsu, D.T., Jorde, U.P., 3D Printing to Guide Ventricular Assist Device Placement in Adults With Congenital Heart Disease and Heart Failure. JACC: Heart Failure 4, 301–311. (2016). https://doi.org/10.1016/j.jchf.2016.01.012

[6] McGovern, E.M., Tretter, J.T., Moore, R.A., Goldstein, B.H., Simulation to Success. JACC: Case Reports 2, 486–487. (2020). https://doi.org/10.1016/j.jaccas.2019.11.078

[7] Crafts, T.D., Ellsperman, S.E., Wannemuehler, T.J., Bellicchi, T.D., Shipchandler, T.Z., Mantravadi, A.V., Three-Dimensional Printing and Its Applications in Otorhinolaryngology–Head and Neck Surgery. Otolaryngol.--head neck surg. 156, 999–1010. (2017). https://doi.org/10.1177/0194599816678372

[8] Wong, R.H.L., Lee, A.P.-W., Three-dimensional printing in structural heart disease and intervention. Ann. Transl. Med. 7, 579–579. (2019). https://doi.org/10.21037/atm.2019.09.73

[9] Lau, I., Sun, Z., Three-dimensional printing in congenital heart disease: A systematic review. J of Medical Radiation Sci 65, 226–236. (2018). https://doi.org/10.1002/jmrs.268

[10] Capan, M., Khojandi, A., Denton, B.T., Williams, K.D., Ayer, T., Chhatwal, J., Schwartz, J.S., From Data to Improved Decisions: Operations Research in Healthcare Delivery. Medical Decision

Making, (2017). https://doi.org/10.1177/0272989x17705636

[11] Burke, E.K., De Causmaecker, P., Berghe, G.V., Van Landeghem, H., The State of the Art of Nurse Rostering. Journal of Scheduling, (2004). https://doi.org/10.1023/b:josh.0000046076.759

[12] Green, L.V., Savin, S., Reducing delays for medical appointments: A queueing approach. Operations Research, 56(6), 1526–1538. (2008). https://doi.org/10.1287/opre.1080.0575

[13] Gupta, D., Denton, B., Appointment scheduling in health care: Challenges and opportunities. IIE Transactions, (2008). https://doi.org/10.1080/07408170802165880

Solving Unrelated parallel machine scheduling problem by genetic algorithm

Hind Bouifalioune[1], Majda Fikri[2], Ghizlane Bencheikh [3]

[1,3] Moulay Ismaïl University of Meknes, Morocco
[2] Ibn Zohr University of Agadir, Morocco
(E-mail : h.bouifalioune@edu.umi.ac.ma ; majdafikri@yahoo.fr ; g.bencheikh@umi.ac.ma)

Abstract

This study addresses unrelated parallel machine scheduling problem (UPMSP), focusing on minimizing the maximum completion times (makespan). To tackle this NP-hard problem, we propose a genetic algorithm (GA) implemented in Python. The efficiency of the GA in exploring and exploiting the search space makes it a suitable choice for scheduling problems characterized by complexity and multiple constraints. Our implementation integrates adapted genetic operators and a fitness function designed to optimize the makespan specifically for UPMSP. Experimental results demonstrate the algorithm's ability to produce high-quality solutions, highlighting its potential as a robust tool for scheduling in industrial applications.

Keywords: *Metaheuristics, Genetic Algorithm, Scheduling Problrm, Python, Combinatorial Optimization, Parallel Machine Scheduling, Operations Research, Optimization.*

1 Introduction

In recent manufacturing industries, increasing productivity and minimizing production costs are essential for a sustainable business. Therefore, scheduling is becoming increasingly important to minimize production costs by reducing production completion times and efficiently allocating resources. Effective scheduling ensures that machines and labor are optimized, reducing downtime and improving workflow. By carefully planning the production process, companies can meet deadlines, improve product quality, and respond more flexibly to market demands while controlling expenses. Consequently, advanced scheduling techniques and tools are essential to maintain competitiveness in the modern manufacturing landscape.

Moreover, the manufacturing process relies on the characteristics of parallel machines. A classic parallel machine system (PMS) can be categorized as identical, uniform, or unrelated:

- **Identical parallel machines:** In this category, all machines are identical in terms of processing capabilities. Each machine can perform the same tasks at the same speed, which simplifies the scheduling process since any job can be assigned to any machine without affecting the overall processing time.
- **Uniform parallel machines:** In a uniform PMS, the machines differ in their processing speeds. Each machine performs tasks at different rates, but the speed ratio remains constant for different jobs. This means that although the machines are not identical, their performances are predictable and can be adjusted based on their known speeds.
- **Unrelated parallel machines:** This category includes machines that vary both in terms of processing capabilities and efficiency according to the jobs. The performance of each machine can differ depending on the specific task, making scheduling more complex. Job allocation must take these variations into account to optimize overall productivity and minimize completion times.

In this study, we aim to solve the problem of unrelated parallel machine scheduling using Genetic Algorithm metaheuristic. Genetic algorithms are a powerful optimization technique inspired by the process of natural selection and evolution. By mimicking genetic processes such as selection, crossover, and mutation, these algorithms can efficiently explore the solution space and find near-optimal solutions for complex scheduling problems.

The implementation of the genetic algorithm involves representing potential solutions as individuals in a population, with each individual representing a possible scheduling arrangement for the parallel machines. These individuals are then evaluated based on objective criteria such as minimizing the makespan or total completion time. Through the iterative process of selection, crossover (recombination), and mutation, the algorithm evolves the population towards better solutions over successive generations.

In our Python implementation, we define appropriate data structures to represent the scheduling problem, including job characteristics, machine assignments, and scheduling constraints. We will then develop functions to initialize the population, evaluate individuals, perform selection, crossover, and mutation operations, and finally, execute the evolutionary process to find an optimal or near-optimal solution.

By leveraging Python's flexibility and extensibility, as well as its rich ecosystem of libraries for scientific computing and optimization, we can create a robust and efficient implementation of the genetic algorithm suited to solving the problem of unrelated parallel machine scheduling. This approach offers the potential to find high-quality solutions within reasonable computational times, contributing to advances in scheduling optimization in manufacturing and related fields.

This paper is organized as follows: in Section 2, we present a brief review of literature dealing with this problem. In Section 3, mathematical model of the problem is described. Section 4, provides a detailed description of the proposed genetic algorithm. In Section 5, experimental results of the proposed approach are presented and a comparison with the Cplex results is discussed. Finally, conclusions and future research directions are provided in Section 6.

2 Literature review

In the literature, we can find several heuristic and metaheuristic algorithms for the mentioned problem. However, most of them are centered on the case of identical parallel machines. In [1], a heuristic is proposed for the case of identical parallel machines with sequence-dependent setup times, aiming to minimize the makespan. A tabu search algorithm is presented [2] with the objective of minimizing the total completion time. A three-phase heuristic is proposed [3] for the same problem with sequence-dependent setup times (independent of the machine) and aiming to minimize the sum of weighted tardiness.

In [4], the authors propose several heuristics and a genetic algorithm to minimize the makespan. Other heuristics for the same problem are proposed by [5,6]. In both cases, the objective is to minimize the makespan, and in the latter case, precedence constraints are also considered. In [7,8], heuristics are proposed for family setup times. [9] propose a linear programming approach where task splitting is also considered.

[10,11] respectively present heuristic and metaheuristic methods for the same problem. The case of unrelated parallel machines with sequence-dependent setup times has been less studied, and only a few articles can be found in the literature. A tabu search algorithm is presented in [12] with the objective of weighted tardiness. Another heuristic for the case of unrelated parallel machines aiming to minimize the weighted mean completion time is proposed [13].

[14] proppposed a simulated annealing method aiming to minimize total tardiness. In [15,16] ,a heuristic and a tabu search algorithm were proposed with the objective of minimizing total weighted tardiness

and maximum tardiness, respectively. The same problem is also studied in [17,18,19], where resource constraints are also considered, with the objective of minimizing makespan, maximum tardiness, and total tardiness, respectively. In [20], a heuristic for the case of unrelated machines aiming to minimize the makespan is also presented.

[21] proposed a method based on the VNS (Variable Neighborhood Search) strategy for both cases, identical and unrelated parallel machines, with the objective of minimizing the makespan. In [22,23], the authors proposed a simulated annealing method and a GRASP (Greedy Randomized Adaptive Search Procedure) algorithm with the objective of minimizing total flow time and total tardiness, respectively.

Regarding exact methods, there are a few articles available in the literature for the parallel machine problem. However, most of them can only solve instances with a small number of tasks and machines (more details in [24,25]).

In this article, we address the problem of unrelated parallel machine scheduling in which setup times depend on the assigned machine. We evaluate and compare some of the methods available in the literature. We also propose a genetic algorithm that shows excellent performance for a large set of benchmarks.

3 Problem's description :

In this paper, we proposed to solve the unrelated parallel machines problem which belongs to the NP-hard class of optimization problems, the problem is stated as a triplet: $P_m \mid S_{i,j,k} \mid C_{max}$, where :

- **P_m:** describes the environment of unrelated parallel machines. The reference to unrelated machines means that the processing time of a job P_{ki}, that is, the processing time of job i on machine k, depends on the machine to which it is assigned, and there is no relationship between the speeds of the machines.
- **$S_{i,j,k}$** :describes the scheduling constraints, which in this case are sequence-dependent setup times. The setup time required for job i when it follows job j on machine k is likely different from that required to prepare job i if it precedes job j on the same machine, such that $S_{i,j,k} \neq S_{j,i,k}$ for each $i, j \in \{0,1,\ldots,n\}$.
- **C_{max} :** is the objective function, which is the makespan.

The scheduling model discussed here deals with the organization of collection of jobs to be executed efficiently on several independent machines. It can be described as follows: we have a set of n independent jobs, $i = \{1,2, \ldots, n\}$ with positive processing times P_i, to be processed on a set of unrelated parallel machines $M = \{1,2,\ldots,m\}$. All jobs are available for processing at time zero, and once a job starts being processed on one of the machines, it must be completed. In other words, job preemption is not allowed. The objective of this paper is therefore to find a non-preemptive schedule $X = \{X_1, X_2, \ldots, X_n\}$ on a set of unrelated parallel machines k ($k = 1,2,\ldots, m$) for jobs j ($j = 1,2,\ldots,n$), in order to minimize the makespan. The diagram in Figure 1 provides a visual representation of this scenario.

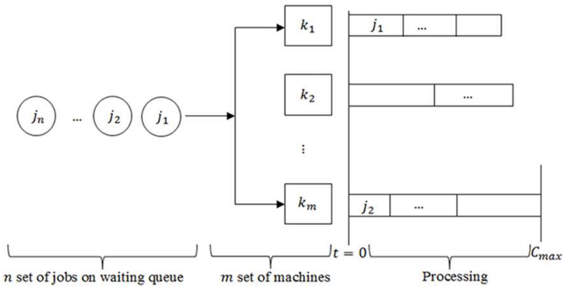


FIG .1 – Parallel machine scheduling problem model [26].

3.1. Mathematical model

In this section, we present a Mathematical Model for the problem that we try to solve. The parallel machine scheduling problem depends on the sequence setup times of each machine by all the jobs assigned to it, this problem is focuses to minimize of the maximum completion time of the schedule. therefore, we used the model proposed by [25].

Notations

Paramètre :

- N : set $\{1, \dots, n\}$ of n jobs.
- M : set $\{1, \dots, m\}$ of m machines.
- p_{ki} : Processing time of job i on machine k.
- V : Sufficiently large number

Decision variables

- $x_{kji} \begin{cases} 1, & \text{if job } i \text{ precedes job } j \text{ on machine } k. \\ 0, & \text{otherwise} \end{cases}$
- C_{kj} **:** Completion times of job j at machine k.
- C_{max} : Makespan, or maximum completion time.

Mathematical Model

Objective function is :

$$min\ C_{max} \qquad (1)$$

$$\sum_{k\in M} \sum_{\substack{j\in\{0\} \cup N \\ j\neq i}} x_{kji} = 1, \forall\, i \in N \qquad (2)$$

$$\sum_{k\in M} \sum_{\substack{i\in N \\ j\neq i}} x_{kji} \leq 1\,, \forall\, j \in N \qquad (3)$$

$$\sum_{i\in N} x_{k0i} \leq 1\,, \forall\, k \in M \qquad (4)$$

$$\sum_{\substack{h\in\{0\}\cup N \\ h\neq i, h\neq j}} x_{khj} \geq x_{kji}\,, \forall j, i \in N, j \neq i, \forall k \in M \qquad (5)$$

$$C_{ki} + V\big(1 - x_{kji}\big) \geq C_{kj} + p_{ki},\ \forall\, j \in \{0\} \cup N, \forall i \in N, j \neq i, \forall k \in M \qquad (6)$$

$$C_{k0} = 0, \forall\, k \in M \qquad (7)$$

$$C_{kj} \geq 0, \forall j \in N, \forall\, k \in M \qquad (8)$$

$$C_{max} \geq C_{kj}\,, \forall j \in N, \forall k \in M \qquad (9)$$

$$x_{kji} \in \{0,1\}, \forall j \in \{0\} \cup N, \forall i \in N, j \neq i, \forall\, k \in M \qquad (10)$$

The goal is to minimize $\boldsymbol{C_{max}}$, which likely represents the makespan or the maximum time required to complete all tasks. The constraint (2) ensures that each task $\boldsymbol{i}$ is assigned exactly once to a machine $\boldsymbol{k}$. Here, $\boldsymbol{x_{kji}}$ is a binary variable that determines whether task $\boldsymbol{i}$ is assigned to machine $\boldsymbol{k}$ and the predecessor to task $\boldsymbol{j}$. The sum ensures that each task is assigned exactly once, avoiding duplication across machines or jobs. The constraint (3) ensures that each job j can be processed by at most one machine $\boldsymbol{k}$ and at most one other task $\boldsymbol{i}$, avoiding overloading a machine with too many jobs. The constraint (4) limits the number of tasks starting from the initial position 0 for each machine $\boldsymbol{k}$. Essentially, this constraint ensures that each machine can have at most one job assigned to the initial slot 0. The constraint (5) ensures that if job $\boldsymbol{i}$ is assigned to machine $\boldsymbol{k}$, some job $\boldsymbol{h}$ (from the set of jobs or the initial position) precedes job $\boldsymbol{i}$. It guarantees the correct ordering of tasks in the schedule. The constraint (6) ensures that if job $\boldsymbol{i}$ precedes job $\boldsymbol{j}$ on machine $\boldsymbol{k}$, the completion time of job $\boldsymbol{i}$ $(\boldsymbol{c_{ki}})$ plus its processing time $(\boldsymbol{p_{ki}})$ must be less than or equal to the start time of job $\boldsymbol{j}$ $\big(\boldsymbol{c_{kj}}\big)$. The term $\boldsymbol{V(1 - x_{kji})}$ adds a large number V to deactivate the constraint when $\boldsymbol{x_{kji} = 0}$, meaning job $\boldsymbol{i}$ does not precede $\boldsymbol{j}$. The constraint (7) initializes the start time for the machine $\boldsymbol{k}$ to 0, assuming that machines start processing jobs at time 0. The constraint (8) ensures that the completion time for every job $\boldsymbol{j}$ on every machine $\boldsymbol{k}$ is non-negative. the constraint (9) ensures that the makespan ($\boldsymbol{c_{max}}$) is greater than or equal to the completion time of any job $\boldsymbol{j}$ on machine $\boldsymbol{k}$. Essentially, it sets the makespan as the maximum of all completion times. Finally, the constraint (10) defines the $\boldsymbol{x_{kji}}$ as binary takes values of either 0 or 1, indicating whether or not job $\boldsymbol{i}$ is assigned to machine $\boldsymbol{k}$ and predecessor job j.

4 The proposed genetic algorithm

The **genetic algorithm** is a widely recognized methodology in the field of metaheuristics, designed to converge toward optimal solutions. This type of algorithm is frequently employed in solving combinatorial optimization problems, particularly in the context of scheduling problems. Figure 2 provides the flowchart of the standard GA.

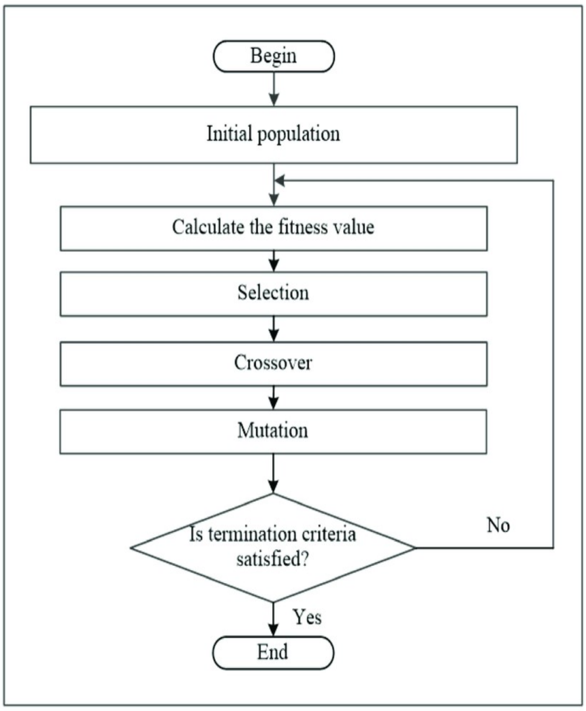


FIG.2 – Flowchart of the standard genetic algorithm [27].

4.1. Chromosome Representation

The first step in a Genetic Algorithm (GA) is the representation of solutions. While several representation methods have been proposed for scheduling problems [28, 29, 30, 31, 32], this paper adopts the Machine List Encoding (MLE) representation of solutions [25]. This encoding is implemented as a dictionary in Python, an efficient method for organizing and managing data in key-value pairs. In Python, dictionaries are built-in data structures that allow fast association of unique keys with corresponding values. In this representation, machines serve as keys, and lists of tasks serve as values within an individual.

This choice is justified not only for its simplicity but also because, in preliminary experiments, it has proven to be the most effective among the various representations tested. Figure 3 shows a solution using MLE for a scheduling problem with three machines and nine tasks. In this encoding, each machine is assigned a list of tasks to be performed in a specific order. For example, tasks 3 and 7 are assigned to machine 0, tasks 5, 1, 2, and 0 to machine 1, and tasks 6, 4, and 8 to machine 2.

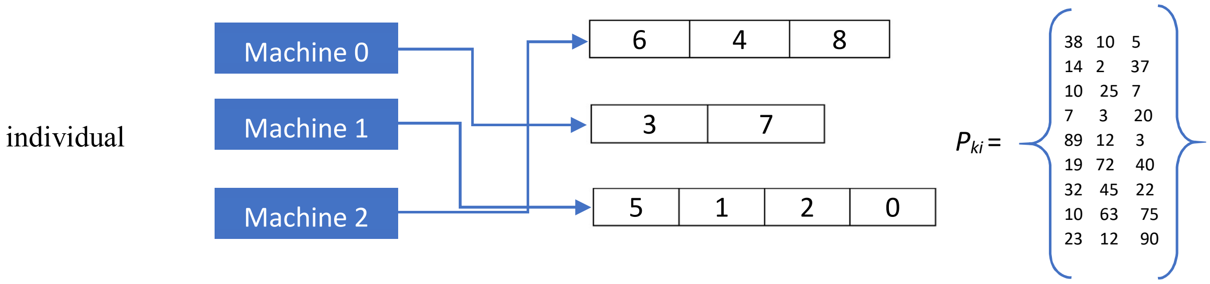


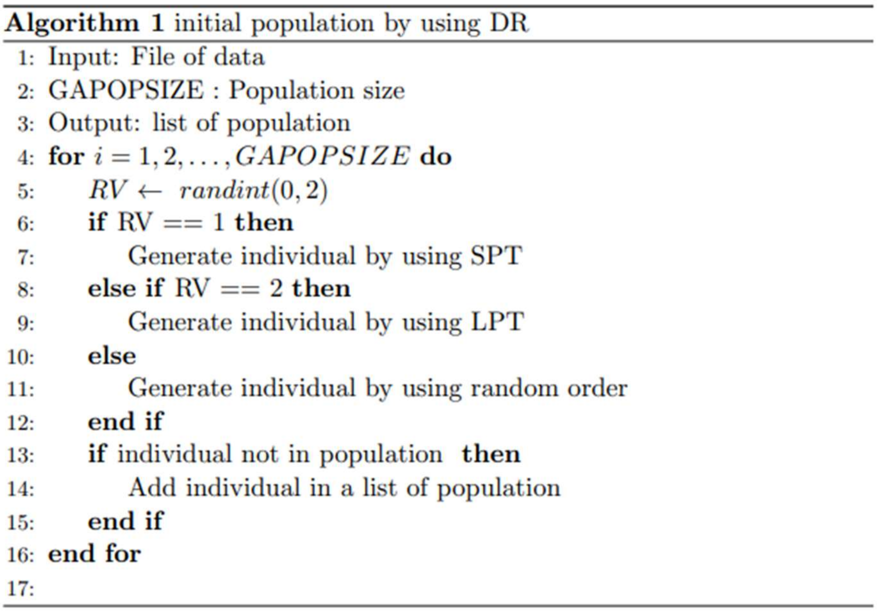
FIG.3 – A solution represented by the machine list encoding

4.2. Generalization of initial Population

The initialization of the population of individuals in a genetic algorithm is crucial for searching efficient solutions. To achieve this, we have chosen to use dispatching rules that define the order of task assignment to machines [35]. Three dispatching rules have been selected:

- **Shortest Processing Time (SPT)** rule prioritizes tasks with the shortest processing time on each machine. This method aims to minimize the total processing time by optimizing the use of available resources.
- **Longest Processing Time (LPT)** rule gives priority to tasks with the longest processing time, which can help in better workload distribution and balancing production across machines.
- **Random Order method** assigns tasks randomly, providing a more diversified approach that can explore different configurations in a non-deterministic manner. While this approach may require more iterations to converge to optimal solutions, it can be useful for exhaustively exploring the solution space.

In the process of initializing the population of individuals, Algorithm 1 is used to randomly select one of these generation methods. This approach ensures diversity in initial solutions, which is essential for effectively exploring the space of potential solutions.



Algorithm 1 initial population by using DR

```
Input: File of data
GAPOPSIZE : Population size
Output: list of population
for $i = 1, 2, \ldots, GAPOPSIZE$ do
    $RV \leftarrow randint(0, 2)$
    if RV == 1 then
        Generate individual by using SPT
    else if RV == 2 then
        Generate individual by using LPT
    else
        Generate individual by using random order
    end if
    if individual not in population then
        Add individual in a list of population
    end if
end for
```

4.3. Individual's evaluation

In this step of the GA, calculate the fitness value for each individual in the population. This value is determined using a specific formula (11) based on the inverse of the maximum sum of processing times of tasks assigned to the machine.

$$fitness = 1/\sum_{k\in M}\sum_{i\in J} p_{ki} \tag{11}$$

To illustrate this, let's take the example of a specific individual as shown in Figure 3. Applying the formula, the maximum sum of processing times of tasks assigned to each machine for this individual is 241. Next, we take the inverse of this sum to obtain the fitness value of this individual, which is 0.00414. This fitness calculation approach is essential as it allows evaluating the quality of each individual based on its ability to minimize the total processing time of tasks on the machines. Thus, individuals that efficiently handle tasks will have a higher fitness value, while those requiring more time will have a lower fitness value.

This process then guides the selection of the most performing individuals for reproduction based on techniques inspired by evolutionary biology, including crossover and mutation. In this framework, individuals demonstrating the best performance according to predefined criteria are chosen to reproduce, combining their characteristics in the hope of creating even more performing offspring.

4.4. Crossover

The crossover operator used in this context is called single-point crossover. As illustrated in Figure 4, for each machine permutation list, a random crossover point is chosen. The child individual is created by directly copying all tasks that appear before the crossover point in the first parent. To ensure that the lists in the child individual represent valid permutations, any tasks after the crossover point from the second parent are not copied directly. Instead, each task in the permutation list of the second parent is checked to see if it already exists in any of the machine lists of the child individual. If a task is already present, it is ignored; otherwise, it is added to the corresponding machine in the child individual. This process ensures that the child individual represents a valid solution [35].

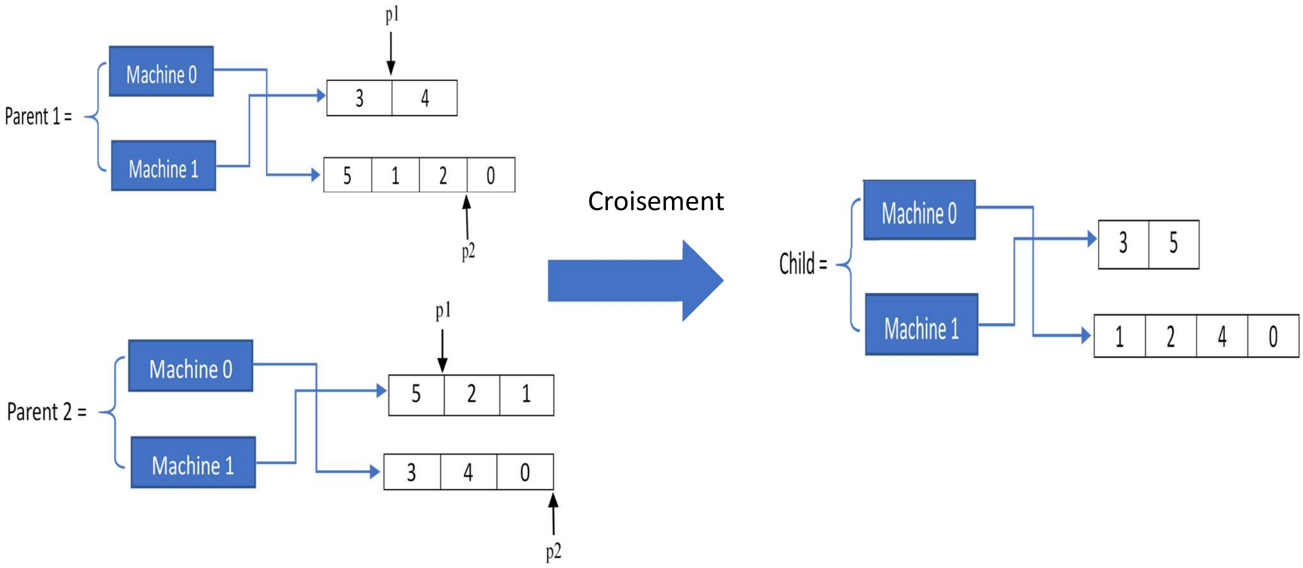


FIG.4 – Point crossover used for MLE

In the example from Figure 4, the crossover point for the first machine was chosen so that the first task is taken from the first parent, while for the second machine, only the first three tasks would be taken from the first parent. Next, the list for the first machine from the second parent is traversed, during which task 5 is added to the child's permutation list, while tasks 2 and 1 are ignored because they are already present in the second machine's list. The same procedure is also used to determine which tasks will be placed on the second machine of the child individual.

4.5. Mutation

The mutation operation, the goal is to modify the properties of an individual to introduce genetic diversity into the population. The mutation operation we use is applied to one of the children. This described operation involves randomly selecting two tasks in a machine list and comparing their respective execution times. If the execution time of the first task is greater than that of the second, their positions in the list are exchanged. In other words, task 2 is placed in the position of task 1, and vice versa. This procedure allows for reorganizing tasks to optimize their order based on their execution times, as illustrated in Figure 5.

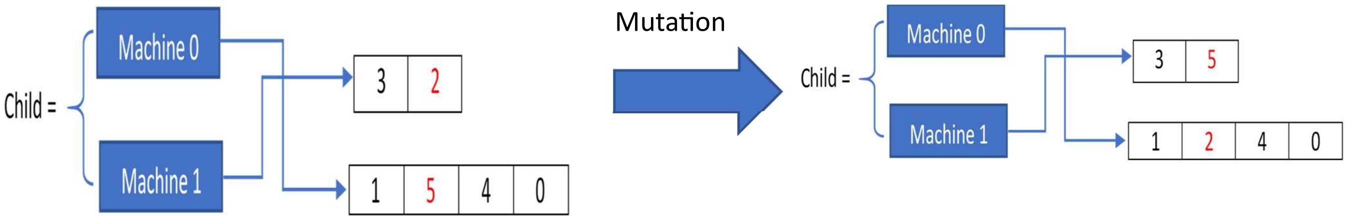
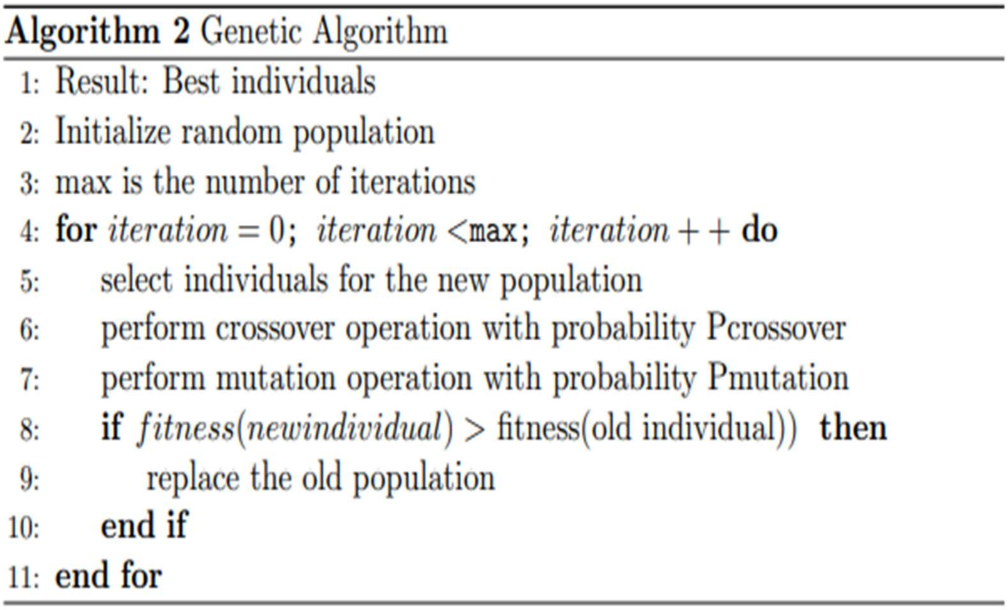


FIG.5 – Description of the mutation method

The idea of this method is to introduce modifications in individuals to maintain genetic diversity within the population and avoid falling into local minima. This allows for exploring new potential solutions and improving the overall quality of solutions in successive generations of the genetic algorithm.

4.6. Genetic Algorithm

After the processes of selection, crossover, and mutation are completed, the entire population of offspring C is created and transferred to the next population P. P is then used in the next iteration, which restarts the entire process. Iterations will stop if there is convergence of results or if the maximum number of iterations is exceeded. The computation procedures for the genetic algorithm are illustrated as shown in the algorithm 2.



Algorithm 2 Genetic Algorithm

```
Result: Best individuals
Initialize random population
max is the number of iterations
for $iteration = 0$; $iteration <$ max; $iteration ++$ do
    select individuals for the new population
    perform crossover operation with probability Pcrossover
    perform mutation operation with probability Pmutation
    if $fitness(newindividual) >$ fitness(old individual)) then
        replace the old population
    end if
end for
```

5 Experimental results

In this section, we present the results of the Unrelated Parallel Machines Scheduling Problem (UPMSP) described in Section 3. We begin by searching for optimal parameters for our algorithm to achieve high performance. Subsequently, we focus on solving various small and large sized test instances. Additionally, the GA algorithm we propose is implemented on an Intel Core i7 processor, operating at 2.07 GHz speed and equipped with 12 GB of RAM.

5.1. Instance Characterization

To conduct computational experiments, we utilized benchmark instances inspired by [25]. In these instances, job processing times are randomly generated following a uniform distribution between 1 and 99. Two categories of problem sizes were considered. The first category, referred to as small instances, includes combinations of jobs (n) and machines (m) with $n = \{6,8,10\}$ and $m = \{2,3,4,5\}$. The second category, large instances, involves larger-scale problems with $n = \{50,100,150,200,250\}$ and $m = \{10,15,20,25,30\}$. This approach ensures a diverse and comprehensive set of test cases to evaluate the algorithm's performance across both small and large problem scenarios.

5.2. The parameters of the proposed genetic algorithm

The first step in optimizing our genetic algorithm (GA) was to fine-tune several critical parameters : crossover probability, mutation probability, and initial population size. These parameters play a vital role in the algorithm's performance. Crossover probability determines how traits from parent solutions are combined to generate new solutions, with higher probabilities encouraging greater exploration of the solution space. Mutation probability controls how frequently modifications are introduced into individual solutions, which helps maintain diversity and avoid premature convergence. The initial population size is equally important, as it impacts the genetic diversity and the algorithm's ability to effectively explore the search space. Striking a balance between these parameters is key to improving the GA's efficiency and effectiveness.

To evaluate their impact, the algorithm was systematically tested using a scenario with two machines and 10 jobs. Table 1 summarizes the results of these tests, which included varying the population size across 50, 100, and 200; the crossover probability at 0.75, 0.8, and 0.9; and the mutation probability at 0.25, 0.15, and 0.05. Each combination was likely run multiple times to ensure reliability, with average execution times calculated to account for variability.

Initial population size	**Probability of crossing**	**Probability of mutation**	**Average execution time (s)**
50	0.75	0.25	0.984
50	0.8	0.15	1.428
50	0.9	0.05	2.090
100	0.75	0.25	0.597
100	0.8	0.15	0.491
100	**0.9**	**0.05**	**0.210**
200	0.75	0.25	0.283
200	0.8	0.15	0.427
200	0.9	0.05	0.534

TAB. 1 – The results obtained from different settings of genetic algorithm parameters

The table demonstrates clear trade-offs, such as shorter execution times with smaller population sizes or specific mutation probabilities. This thorough parameter tuning process identified the best-performing configuration for optimizing the algorithm, which is highlighted in bold in the table. These results provide practical guidance for selecting GA parameters that achieve a balance between computational efficiency and solution quality.

5.3. Comparison with Cplex solver for small instances

After configuring the parameters, the metaheuristic Genetic Algorithm (GA) was applied to small instances to evaluate its performance against the CPLEX solver for the Unrelated Parallel Machine Scheduling Problem (UPMSP) discussed in Section 3. These small instances, defined by the number of jobs and machines, determined the dimensions and complexity of the problems. Table 2 presents the results of this comparison, highlighting makespan (C_{max}) values and solution times for each test case. Both methods achieved identical makespan values for most problems, demonstrating that the GA is competitive in terms of solution quality. However, the execution times reveal key differences: while CPLEX is generally faster for smaller problems, such as problems 1 and 3, its performance deteriorates as complexity increases. This is evident in problem 9, where its execution time rises sharply to 136.98 seconds compared to the GA's 0.14 seconds.

Problem	Job	Machine	CPLEX		GA	
			C_{max}	Execution time (s)	C_{max}	Execution time (s)
1	**6**	2	83	0.28	83	0.16
2		3	62	0.55	62	0.70
3		4	59	0.11	59	0.15
4		5	40	0.19	40	16.14
5	**8**	2	117	0.53	117	0.32
6		3	80	0.48	80	0.58
7		4	52	0.64	52	2.03
8		5	50	0.17	50	1.49
9	**10**	2	203	136.98	203	0.14
10		3	101	2.03	101	3.01
11		4	77	3.05	77	62.62
12		5	-	-	68	21.15

TAB. 2 – The performance of the genetic algorithm for small instances.

In contrast, the GA demonstrates consistent execution times and excels in solving larger and more complex instances. For example, in problem 12, the GA successfully provides a solution when CPLEX fails to solve the largest instance (10 jobs and 5 machines). The GA's ability to solve all instances, including the most complex cases, with competitive or superior execution times underscores its robustness and efficiency. These findings confirm that the GA is a reliable and scalable alternative, capable of delivering optimal or near-optimal solutions efficiently, even for larger and more computationally challenging problems.

5.4. Comparison with existing algorithm

Since the CPLEX solver cannot solve medium and larger instances to optimality, and its obtained lower bounds are insufficient for verification, validating the performance of the Genetic Algorithm (GA) for these instances requires comparison with other algorithms. To address this, we compare our GA with two variants of GA proposed in article [25], which we refer to as GAV1 and GAV2. The results are reported in terms of the makespan (C_{max}) for each algorithm, with performance evaluated after running five replicates. This approach helps ensure the robustness of the findings by accounting for the inherent variability of the algorithms. By repeating the experiments multiple times, the impact of stochastic factors is reduced, allowing for a more accurate and dependable comparison of the GA's performance against other algorithms. This method provides a more reliable assessment of each algorithm's effectiveness, offering a clearer picture of their true capabilities.

As shown in Table 3, the results for the large instances are summarized by averaging the 40 instances in each $n \times m$ group, providing a thorough comparison. The worst results for each $n \times m$ set are indicated in italics, while the best results are highlighted in boldface. This format allows for a clear visualization of the performance extremes, offering insights into both the algorithm's best capabilities and its limitations across the different problem configurations.

Job	Machine	GA	GAV1	GAV2
50	10	**200**	*210*	**200**
	15	**156**	*169*	162
	20	**138**	*173*	147
	25	**116**	*135*	*135*
	30	**114**	*131*	129
100	10	427	*437*	**414**
	15	*342*	*342*	**341**
	20	**277**	*315*	286
	25	**236**	*290*	236
	30	**198**	*247*	**198**
150	10	**686**	*718*	678
	15	**508**	*544*	511
	20	**385**	*399*	**385**
	25	**343**	*370*	**343**
	30	**278**	*337*	288
200	10	931	*975*	**911**
	15	**655**	*709*	674
	20	516	*564*	**509**
	25	**446**	452	*488*
	30	**381**	*414*	394
250	10	1193	*1202*	**1157**
	15	824	*906*	**817**
	20	**684**	*757*	696
	25	**604**	*629*	**604**
	30	**454**	*502*	481

TAB. 3 – The performance of the genetic algorithm for large instances.

Our implementation of the Genetic Algorithm (GA) in Python leverages the efficiency of dictionaries (acting like linked lists) and carefully designed crossover and mutation methods to achieve superior results quickly, even for larger problem instances. By using dictionaries, we efficiently manage and access population data, enabling fast searches and updates that significantly accelerate the algorithm's execution. The crossover method is designed to intelligently combine parent solutions, preserving their most advantageous traits while generating diverse offspring that explore new areas of the solution space. Our mutation method introduces variability in a controlled manner, ensuring the algorithm does not get stuck in local optima and continuously improves solution quality. Together, these methods enhance the GA's ability to quickly find high-quality solutions. Furthermore, our approach scales well to larger instances, maintaining performance and accuracy, thus demonstrating its robustness and adaptability to complex real-world problems.

6 Conclusion

In conclusion, this article addressed the resolution of the Unrelated Parallel Machine Scheduling Problem (UPMSP) in multi-product and multi-machine workshops using a Genetic Algorithm (GA). We demonstrated that employing crossover, mutation operations, and population initialization based on priority rules, combined with a dictionary-based data structure in Python, efficiently optimizes production scheduling.

The results obtained show a significant reduction in production tardiness, highlighting the effectiveness of this metaheuristic approach for UPMSP. However, the potential of this method does not end there. Future research should explore integrating the production scheduling problem (UPMSP) with the Vehicle Routing Problem (VRP) for product delivery.

This combination would minimize not only production tardiness but also product delivery delays, leveraging advanced metaheuristic methods and machine learning techniques. Thus, an integrated approach could offer even more robust and efficient solutions to optimize the entire production and distribution process, meeting the increasing demands for flexibility and efficiency in modern supply chains.

In conclusion, the development and application of hybrid algorithms combining GA, VRP, and machine learning promise new advances in production and logistics optimization, contributing to enhanced competitiveness and sustainability of businesses.

References:

[1] Guinet, A., 1993. Scheduling sequence-dependent jobs on identical parallel machines to minimize completion time criteria. *International Journal of Production Research 31*, 1579–1594.

[2] Franca, P., Gendreau, M., Laporte, G., Mnller, F., 1996. A tabu search heuristic for the multiprocessor scheduling problem with sequence dependent setup times. *International Journal of Production Economics 43*, 79–89.

[3] Lee, Y., Pinedo, M., 1997. Scheduling jobs on parallel machines with sequence dependent setup times. *European Journal of Operational Research 100*, 464– 474.

[4] Kurz, M., Askin, R., 2001. Heuristic scheduling of parallel machines with sequencedependent set-up times. *International Journal of Production Research 39*, 3747– 3769.

[5] Gendreau, M., Laporte, G., Morais-Guimaraes, E., 2001. A divide and merge heuristic for the multiprocessor scheduling problem with sequence dependent setup times. *European Journal of Operational Research 133*, 183–189.

[6] Hurink, J., Knust, S., 2001. List scheduling in a parallel machine environment with precedence constraints and setup times. *Operations Research Letters 29*, 231–239.

[7] Eom, D., Shin, H., Kwun, I., Shim, J., Kim, S., 2002. Scheduling jobs on parallel machines with sequence dependent family set-up times. *International Journal of Advanced Manufacturing Technology 19*, 926–932.

[8] Dunstall, S., Wirth, A., 2005. Heuristic methods for the identical parallel machine flowtime problem with set-up times. *Computers & Operations Research 32*, 2479–2491.

[9] Tahar, D., Yalaoui, F., Chu, C., Amodeo, L., 2006. A linear programming approach for identical parallel machine scheduling with job splitting and sequencedependent setup times. *International Journal of Production Economics 99*, 63–73.

[10] Anghinolfi, D., Paolucci, M., 2007. Parallel machine total tardiness scheduling with a new hybrid metaheuristic approach. *Computers & Operations Research 34*, 3471–3490.

[11] Pfund, M., Fowler, J., Gadkari, A., Chen, Y., 2008. Scheduling jobs on parallel machines with setup times and ready times. *Computers & Industrial Engineering 54*, 764–782.

[12] Logendran, R., McDonell, B., Smucker, B., 2007. Scheduling unrelated parallel machines with sequence-dependent setups. *Computers & Operations Research 34*, 3420–3438.

[13] Weng, M., Lu, J., Ren, H., 2001. Unrelated parallel machine scheduling with setup consideration and a total weighted completion time objective. *International Journal of Production Economics 70*, 215–226.

[14] Kim, D., Kim, K., Jang, W., Chen, F., 2002. Unrelated parallel machine scheduling with setup times using simulated annealing. *Robotics and Computer Integrated Manufacturing 18*, 223–231.

[15] Kim, D., Na, D., Chen, F., 2003. Unrelated parallel machine scheduling with setup times and a total weighted tardiness objective. *Robotics and Computer Integrated Manufacturing 19*, 173–181.

[16] Kim, C., Shin, H., 2003. Scheduling jobs on parallel machines: a restricted tabu search approach. *International Journal of Advanced Manufacturing Technology 22*, 278–287.

[17] Chen, J., 2005. Unrelated parallel machine scheduling with secondary resource constraints. *International Journal of Advanced Manufacturing Technology 26*, 285–292.

[18] Chen, J., 2006. Minimization of maximum tardiness on unrelated parallel machines with process restrictions and setups. *International Journal of Advanced Manufacturing Technology 29*, 557–563

[19] Chen, J., Wu, T., 2006. Total tardiness minimization on unrelated parallel machine scheduling with auxiliary equipment constraints. OMEGA, *The International Journal of Management Science 34*, 81–89.

[20] Rabadi, G., Moraga, R., Salem, A., 2006. Heuristics for the unrelated parallel machine scheduling problem with setup times. *Journal of Intelligent Manufacturing 17*, 85–97.

[21] Rocha de Paula, M., Gómez-Ravetti, M., Robson-Mateus, G., Pardalos, P., 2007. Solving parallel machines scheduling problems with sequence-dependent setup times using variable neighbourhood search. IMA, *Journal of Management Mathematics 18*, 101–115.

[22] Low, C., 2005. Simulated annealing heuristic for flow shop scheduling problems with unrelated parallel machines. *Computers & Operations Research 32*, 2013– 2025.

[23] Armentano, V., Felizardo, M., 2007. Minimizing total tardiness in parallel machine scheduling with setup times: An adaptive memory-based GRASP approach. *European Journal of Operational Research 183*, 100–114.

[24] Allahverdi, A., Ng, C., Cheng, T., Kovalyov, M., 2008. A survey of scheduling problems with setup times or costs. *European Journal of Operational Research 187*, 985– 1032.

[25] Vallada, E., Ruiz, R., 2011. A genetic algorithm for the unrelated parallel machine scheduling problem with sequence dependent setup times. *Eur. J. Oper. Res. 211*, 612–622. https://doi.org/10.1016/j.ejor.2011.01.011.

[26] Ezugwu, A.E., Adeleke, O.J., Viriri, S., 2018. Symbiotic organisms search algorithm for the unrelated parallel machines scheduling with sequence-dependent setup times. *PLOS ONE 13*, e0200030.

[27] Albadr, M.A., Tiun, S., Ayob, M., AL-Dhief, F., 2020. Genetic Algorithm Based on Natural Selection Theory for Optimization Problems. *Symmetry 12*, 1758.

[28] Costa, A., Cappadonna, F. A., & Fichera, S. (2013). A hybrid genetic algo rithm for job sequencing and worker allocation in parallel unrelated machines with sequence-dependent setup times. *The International Journal of Advanced Manufacturing Technology 69*, 2799–2817.

[29] Durasević, M., & Jakobović, D. (2018). Evolving dispatching rules for optimis ing many-objective criteria in the unrelated machines environment. *Genetic Programming and Evolvable Machines 19*, 9–51.

[30] Bean, J. C. (1994). Genetic algorithms and random keys for sequencing and optimization. *ORSA Journal on Computing 6* , 154–160.

[31] Behnamian, J., Zandieh, M., & Fatemi Ghomi, S. (2009). Parallel machine scheduling problems with sequence-dependent setup times using an ACO, SA and VNS hybrid algorithm. *Expert Systems with Applications 36*, 9637–9644.

[32] Balin, S. (2011). Non-identical parallel machine scheduling using genetic algorithm. *Expert Systems with Applications 38*, 6814–6821.

[33] Maheswaran, M., Ali, S., Siegel, H. J., Hensgen, D., & Freund, R. F. (1999). Dynamic Mapping of a Class of Independent Tasks onto Hetero geneous Computing Systems. *Journal of Parallel and Distributed Computing 59*, 107–131.

[34] Braun, T. D., Siegel, H. J., Beck, N., Bölöni, L. L., Maheswaran, M., Reuther, A. I., Robertson, J. P., Theys, M. D., Yao, B., Hensgen, D., & Freund, 30 R. F. (2001). A Comparison of Eleven Static Heuristics for Mapping a Class of Independent Tasks onto Heterogeneous Distributed Computing Systems. *Journal of Parallel and Distributed Computing 61*, 810–837.

[35] Vlašić, I., Đurasević, M., Jakobović, D., 2019. Improving genetic algorithm performance by population initialisation with dispatching rules. *Computers & Industrial Engineering. 137*, 106030.

SUPPORT TOOL TO ASSIST THE CHOICE OF SKILLS FOR MAINTENANCE INTERVENTION BASED ON EXPERIMENTAL DESIGN

. Mariam JARIFI [1], Rachid ABRAIM [2], Badr DAKKAK[3] ,El Hassan IRHIRANE [4]

Mariam JARIFI
Team of Systems Information,
Logistics and Production.
National School of Applied
Sciences
Marrakesh
Morocco
m.jarifi@uca.ac.ma

Rachid ABRAIM
Team of Systems Information,
Logistics and Production.
National School of Applied
Sciences
Marrakesh
Morocco
r.abraim@uca.ac.ma

Badr DAKKAK
Team of Systems Information,
Logistics and Production.
National School of Applied
Sciences
Marrakesh
Morocco
b.dakkak@uca.ac.ma

El Hassan IRHIRANE
Team of Systems Information,
Logistics and Production.
National School of Applied
Sciences
Marrakesh
Morocco
e.irhirane@uca.ac.ma

Abstract : *Industrial maintenance occupies a predominant place in manufacturing companies and it effectively contributes to the improvement and the development of the production tool. However, the major concern of industrialists is to carry out efficient interventions in optimal time without penalizing production. Faced with this challenge, maintenance actors often find it difficult to choose the right skills to assign to the different tasks, depending on the complexity and the rapidity required of each task.*

In this article, we present a tool that helps manufacturers to choose the right actors for maintenance interventions. It is based on the practice of the experimental design and the estimation of the levels of competences of each actor in each maintenance level. Such a tool highlights and takes into account two important responses being the rapidity and quality of interventions. The results obtained were simulated on simulation software and examples of treatment of the results were also presented.

Keywords: *Competences, maintenance levels, experimental design, qualification, decision support tool.*

I. INTRODUCTION

The subject of skills plays a vital role and takes a decisive position in companies, especially in an evolving environment where new forms of organization and modes of productive performance have appeared. Human resources are increasingly managed individually, by taking into account their competence; they are no longer regarded as anonymous resources. It is observed that to perform a given task by two operators of the same qualification, performance varies according to the actor (Dakkak and al. 2014). It brings about the concept of competency level, which is relative to each actor. Many studies focus on the sequencing of activities for human resources, but only few approaches take into account the competency level of these resources (Elizabeth Boye and al. 2015) (Nassazi, 2013) (Boumane, 2007) (Janique Soulié and al. 2003) (Le Boterf and al. 2004). And almost all approaches that are aware to human resources’ skills are mainly the "able or not" type.

The subject of competence management was treated in several research works and with several ways according to the viewing angle of each author using sometimes managerial and empirical approaches and some other times mathematical models.

(Bryan & al. 2002) considers the problem of assigning workers to manufacturing cells in order to maximize the effectiveness of the organization. The authors assumed that the organization effectiveness to be a function of the productivity, output quality, and training costs associated with a particular worker assignment. Traditionally, these worker assignments have been based only on the technical skills of the workers. The proposed model by (Bryan & al. 2002) also includes human skills and permits the ability to change the skill levels of workers by providing them with additional training. The problem is formulated as a mixed integer programming problem. A total of 32 test problems were developed and varied with regard to the total training time, the available training time for each worker, the training costs, the productivity coefficients and the quality level coefficients. Results indicate that this model provides better worker assignments than one that only considers technical skills.

(Imren, 2006) collected the necessary data on the skills required by the engineering and manufacturing employers based on a detailed review of the literature. The author affirmed that it is necessary to understand the needs of the industry to acquire the appropriate engineering skills. This can only be achieved through cooperative, inclusive, transparent and centrally coordinated approaches to skills assessment, monitoring and development.

(Gillian & al. 2007) outline an electronic research skills audit tool that has been developed to map both transferable and discipline-specific skills teaching and assessment within individual modules, the results of which can be collated and analyzed across entire degree programmers'.

(Arsovski et al., 2008) proposes the use of information and communication technologies to improve the quality of the maintenance process.

(Jung, 2015) presented a study on the identification of skills required for geographic information systems (GIS) posts. This study was based on the analysis of 946 GIS announcements from 2007 to 2014. Such analysis permitted to identify the key competences for improving the employability of GIS majors.

(Luis & al. 2015) did a structural equation modelling applied to data collected from 144 manufacturing firms in the automotive industry from several countries. Findings provide evidences that workforce' skills may foster manufacturing flexibility as an effective approach to cape with uncertain environments and turbulent markets. More precisely, results show that employees' skills directly influence new product, volume and mix flexibility, which in turn directly influence business performance.

(MorlyMonteiro& al. 2015) presented a study to elucidate the relationship between risk management and the success of the project given the potential impact of the complexity of the project. The authors have shown that any approach or approach is doomed to failure without taking into account two aspects of skills: soft and hard.

Jorge Moutinho et Fernanda Oliveira (2015) affirm that the logistics and information systems assume relevant rolls to consolidate global performance. Beside efficiency, effectiveness productivity and flexibility, field teams need skills on autonomy responsibility and proactivity.

(Guillermo & al. 2016) affirm that production resources usually represent an important constraint in a manufacturing activity, specially talking about the management of human resources and their skills. They consider an open shop scheduling problem based on a mechanical production workshop to minimize the total flow time including a multi-skill resource constraint. Then, they count with a number of workers that have a versatility to carry out different tasks, and according to their assignment a schedule is generated. In that way, they have formulated the problem a linear as and non-linear mathematical model which applies the classic scheduling constraints, adding some different resources constraints related to personnel staff competences and their availability to execute one task. In addition, they introduce a genetic algorithm and an ant colony optimization (ACO) tool to solve large size problems. The best tool (ACO) has been used to solve a real industrial case.

(Nilesh Pancholi, Mangal Bhatt, 2017) use a multi-attribute decision making (MADM) approach. The primary findings of this research work are to enhance quality in planning the maintenance activities of critical components.

Therefore, the main factors related to the right choice of human resources for a given intervention are the availability of individuals, their number, their suitability for tasks and, finally, their level of competence, which directly influences the performance of a given intervention. The research papers show that the human resource competency levels are poorly integrated into scheduling models. The notion of competence level to quantify the rapidity and performance of a person in relation to a task to be performed is little taken into account.

In this paper, we pay greater attention to the choice of human skills for a given maintenance intervention according to the level of competence of each actor and the complexity of each task. In other words, it is a question of coming up with a tool based on the experimental design to help the industrialists to choose, from the available competences, the actors who might intervene in each maintenance task taking into account both the actors' qualification level, the desired rapidity and quality of intervention.

For this purpose, this paper is organized as follows. First, we begin by definitions of the competence concept. In this section we present the different definitions and the main studies that have been carried out in this field in order to gain a clear understanding of the task. Second, we define the objectives, the hypotheses of the study and the responses to achieve them. Third, we search the factors that could influence the responses, and we define the low and high levels of the factors used. Thus, we choose the right experimental design for each response. Finally, we conclude with a practical study to test the validity of the established models.

II. DEFINITIONS

The competence is the capacity of a person (actor) to act and react with the pertinence required to perform an activity in a work situation. The actor is at the heart of a process, which consists in selecting, combining and mobilizing knowledge, know-how, abilities and behaviors on the one hand and environmental resources on the other, in order to accomplish a mission defined by the organization (Boumane& al., 2006).

On one hand the competence can be considered as a process but also as a disposition to act (Le Boterf, 2004). The competence is a process that builds or adapts action strategies by mobilizing resources to accomplish a given mission. Through this process the subject learns and develops his professional practices. On the other hand, to consider competence solely as a disposition to act is a risk of reducing it to personal resources.

The competence is finalized; it is linked to a mission defined by the work organization. The competence consists in carrying out a mission within the framework of the company's strategy and in the spirit of its culture (Lévy-Leboyer, 1997). Faced with a situation of work, the subject is led to take the initiatives and the decisions necessary to achieve the expected performances. Thus, the competence must be clearly described and formalized in relation to a specific mission.

Classically, the competence is defined as the combined implementation of knowledge, know-how (mastered practices) and know-how (attitudes and behaviors). A competency model proposed by (Harzallah& al. 2006) (figure 1) is based on four characteristics related to the type of competency, skill category, the context in which the competency is used, and the mission associated with that skill.

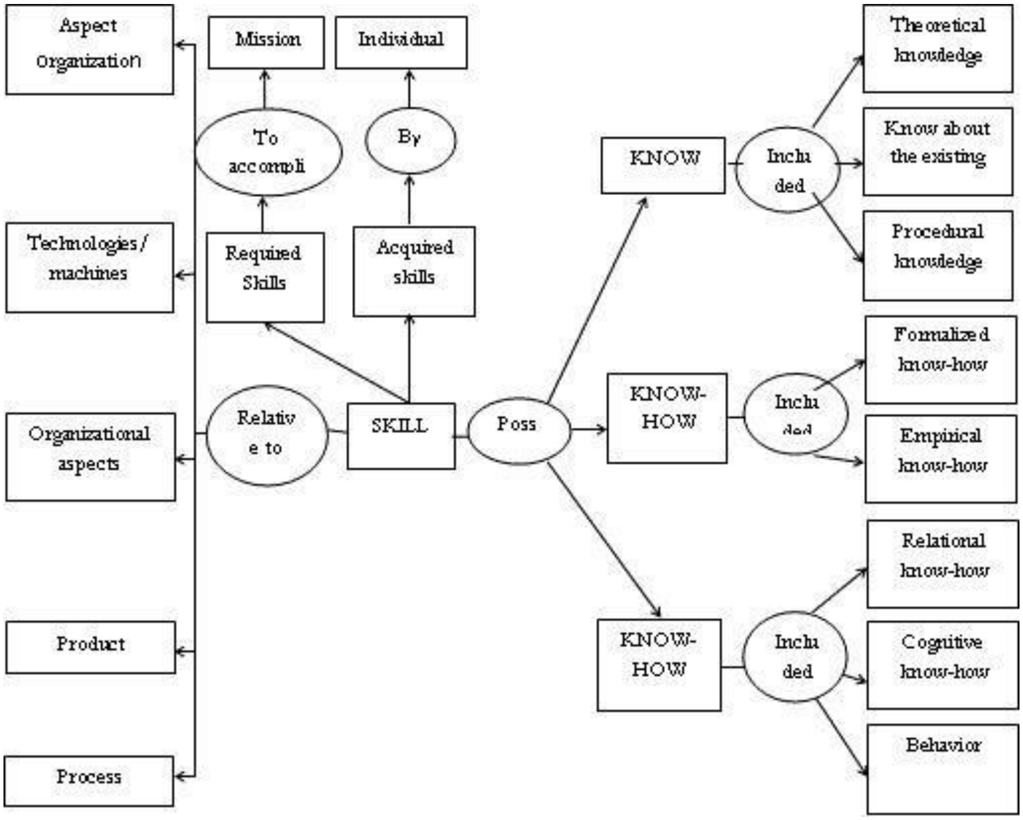


Fig. 1. Skill Skill model (Harzallah et al.2006)

Competence management (CM) refers to how competence is managed with respect to organizations, groups and individuals (Berio&Harzallah 2005). The organizational view has shifted from using the term qualifications to the broader term competence (Boucher & al. 2007). Much research emphasises the organisational perspective related to strategic planning, involving identification of competence gaps and strategies to fill these gaps (Lindgren & al. 2004). They found that for organisational CM activities, skill-based competence descriptions are appropriate and often used. For individual CM activities, job-based descriptions are common. There is a tension between these two views: i) skill-based, as personal descriptions with a present and future focus and; ii) job-based, as positional descriptions with a focus on the past.

III. CONCEPTUAL AND THEORETICAL FRAMEWORK

A. Experimental design

Experimental design is the process of planning a study to meet specified objectives. Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient sample size and power are available to answer the research questions of interest as clearly and efficiently as possible.

Experimental design is a careful balancing of several features including “power”, generalizability, various forms of “validity”, practicality and cost. A thoughtful balancing of these features in advance will result in an experiment with the best chance of providing useful evidence to modify the current state of knowledge in a particular scientific field. On the other hand, it is unfortunate that many experiments are designed with avoidable flaws. It is only rarely in these circumstances that statistical analysis can rescue the experimenter (Howard J. Seltman, 2018).

Experimental design is part of the set of quality tools that allow companies to progress in mastering the design of new products and in mastering manufacturing processes. Together with other statistical tools such as SPC and multiple regression analysis, they form a coherent set of formidable efficiency to solve many quality problems. Ignoring these’s today would deprive oneself of a significant potential for improvement for the company. Whatever the sector of activity and whatever the industry, the latter is always required to conduct tests. But these tests are unfortunately too often conducted without methodology. We proceed by trial and error, without rigorously planning the tests to obtain a plethora of results that we do not always know very well exploit. And yet, the experimental design makes it possible to rigorously conduct the tests with a view to a perfectly defined objective. It will also allow a considerable reduction in the number of tests compared to traditional techniques. More importantly, it will allow a quick and unequivocal interpretation of test results by providing an experimental model of the system under study (Jacques et al 2006).

In fact, this is exactly what the technician is looking for in a problem. The method, once understood, is an irreversible step in the career of the latter who can no longer consider conducting tests without using a plan of experiments.

In this article, we will lead another way of applying the experimental design. It is a question of designing a tool in form of abacuses to help the choice of human skills for maintenance interventions.

B. *Methodology of experimental design*

The different steps of the experimental design are as follows (Guillaume and al, 2005):

1. Clear definition of the problem being studied: proposed targets, consequences of a wrong decision, budget (in time, cost, means, etc.).
2. Compilation of the current local and bibliographical knowledge. If some necessary information is not available, experiments must be undertaken. Therefore, a complete and precise list of the factors likely to be influential, the responses, and the constraints must be established. The area of the experimental domain in which the missing information is to be sought has to be defined. It is referred to as the experimental domain of interest.
3. Setting up an experimental strategy (or experimental design) (i.e., to choose the experiments to be carried out according to the defined targets, the means available, and the desired information). The researcher seeks a relationship of cause and effect between some parameters of the phenomenon (called factors), which are supposed to influence the behavior of the phenomenon and other parameters (called responses) that define the result of the phenomenon. The planning of experiments consists in forcing the factors (input) to vary in a precise way, measuring the induced variations of the answers (output) and then deducing the relationships between causes and consequences.
4. Carrying out the experiments that will give us the values of the studied responses.
5. Deduction of the answers to the questions either directly or with the help of a mathematical model.

IV. EXPERIMENTAL SET-UP AND THE USED METHOD

In this section, we will follow the steps of the experimental design to design a model to help the selection of human skills for carrying out industrial maintenance interventions. To do this, the concepts listed below represent the expected results of the application of the tool:

- Execution of maintenance interventions in optimal time, respecting production requirements and available skills;
- The maintenance intervention must meet the performance requirements imposed in terms of quality, safety and environment. They will be called subsequently the requirement of quality.

The study will be successful if the competences for a given intervention can be identified knowing the desired duration and quality of the intervention.

A. *The responses to achieve the objectives:*

- The first answer matches the first objective. We will therefore choose the rapidity of the intervention.
- The second answer matches the second objective. So we will choose the desired quality as second response.
-

B. *Hypotheses of the study:*

This study will be valid if the following considerations are taken into account:

- For reasons of simplification, we will consider only the first three maintenance levels. However, the study may be generalized later.
- Get the collaboration of the staff so that the work done during the experimentation phase is the true reflection of what really happens
- All members of the stuff are able to perform the three maintenance levels. Only the rapidity and the quality of interventions differ from one actor to another.
- The worksheets are written and formalized.

C. *Investigation of factors expected to affect responses*

What are the factors that may impact the quality and the rapidity of maintenance intervention?

To try not to forget factors, we have to proceed methodically. It is necessary to analyze all the major sectors that may involve factors impacting the chosen responses:

- Maintenance levels;
- Work atmosphere;
- Noise and olfactory disamenities;
- Production pressure;
- Availability of tooling;
- Availability of spare parts;
- Accessibility to place of intervention;
- Season (summer, winter…);
- Day (i.e., the day of the intervention corresponds to an ordinary day, weekend, public holiday, beginning of the week, end of the week…);
- Intervention time (morning, afternoon or night. Because, the performance of operators depends on the intervention time);
- …

To illustrate the approach, we have given a first list of some factors that may have an impact on responses.

From this list, the factors that are expected to be influential are selected and divided into two categories: the factors that will be studied through the experimental design and the factors that will not be studied during the experiment but their levels will always be set at the same value. In this study, three factors were chosen to develop the experimental design: the 1st, the 2nd and the 3rd maintenance level.

D. Levels of input factors

It is a question of choosing the high and the low levels of the three factors selected. To do this, we will link each factor with a closed questionnaire where we will correspond to each question asked a grid of answers with four affirmations: yes, rather yes, rather no and no. Each assertion is associated with a weighting coefficient 1; 0.7; 0.3 and 0 respectively as illustrated in Table 1:

Questions	Affirmation			
	Yes	Rather yes	Rather no	No
Question 1				
...				
Question n				

Table 1. Questionnaire model

The questions of such questionnaire must summarize the tasks and the maintenance interventions to be carried out on the equipment and the means of production. It is then necessary to share them according to the first three maintenance levels.

Once the questionnaire is established, the level of each actor is the sum of the scores obtained in each level of maintenance. Therefore, the high and the low level of each factor can be determined. The low level corresponds to the lowest score and the high level represents the highest score.

It is interesting to summarize these levels in a table (Table 2).

Factor	Low level (-)	High level (+)
1stmaintenance level		
2nd maintenance level		
3rdmaintenance level		

Table 2. Factors and field of study.

E. Choice of the experimental design

In the case of this study, we considered three factors. The low and the high level of each factor have been defined. We assume that the levels of the factors to be kept constant during the experiment are already specified.

Therefore, it is advisable to choose a complete factorial design 2^3:

$$R = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3 \qquad (1)$$

This design can be represented by a figure (Figure 2), indicating the field of study and the points of experiment. The low levels and the high levels of the factors coordinate the points of experience. It can also be represented by the experimental matrix (normal units) or the matrix of experiments (coded units).

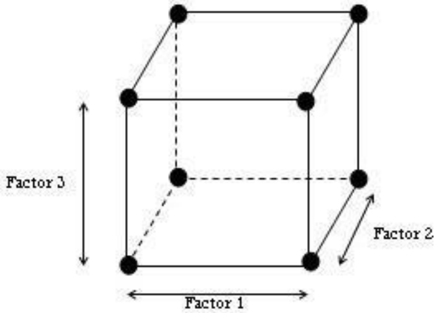


Fig. 2. Representation of experimental design

The experimental matrix (Table 3) includes 6 columns: the names of the trails, the three factors and the two responses (output).

Trials	**Factor 1**	**Factor 2**	**Factor 3**	**Intervention rapidityR_1**	**Intervention quality R_2**
1	**-1**	**-1**	**-1**		
2	**+1**	**-1**	**-1**		
3	**-1**	**+1**	**-1**		
4	**+1**	**+1**	**-1**		
5	**-1**	**-1**	**+1**		
6	**+1**	**-1**	**+1**		
7	**-1**	**+1**	**+1**		
8	**+1**	**+1**	**+1**		
Level -					
Level +					

Table 3. Experimental matrix

The trials must be carried out according to the design of experiments and the results will be recorded in the experimental matrix.

The letters have the following meaning:

- R_i represents the response,
- x_i represents the factor level,
- a_i coefficient represents the factor effect or the first-order effect,
- a_{ij} coefficient represents an interaction between two factors or second-order interaction (binary interaction between xi and x_j),
- a_{123} coefficient represents the interaction between the three factors or third-order interaction (ternary interaction among x_1, x_2 and x_3).

The rapidity and the quality of the interventions are measured as follow:

First, make an inventory of the maintenance activities and tasks that can be performed on the company's machine park;

- Make an estimate of the completion time of each task taking into account all constraints and opportunities. This leads to distinguish three times:

    - ***Optimistic time*** (T_o): if the maintenance task is performed in very good conditions better than expected;
    - ***Realistic time*** (T_r): if the maintenance task is performed under normal conditions;
    - ***Pessimistic time*** (T_p): if the intervention is carried out under bad conditions, worse than expected
- Weigh intervals of interventions as follow (table 4)

Task time	Weight Value
$T \leq T_o$	1
$T_o < T \leq T_r$	2
$T_r < T \leq T_p$	3
$T > T_p$	4

Table 4. Intervention time weighting

- Weigh the quality of interventions as follow (table 5)

Intervention quality	**Weight Value**
Intervention is performed: • With tests, functional checks using additional skills • Without taking into account safety and environmental aspects	1
Intervention is performed: • With tests and functional checks only • Taking into account just some requirements related to safety and environmental aspect	2
The intervention: • Renovates the equipment and it is carried out without any test or verification without calling for additional skills • Takes into account the requirements related to safety and environment.	3

Table 5. Intervention quality weighting

F. Experimentation

Once the high and the low levels of each factor are defined, we move to the experimental phase as shown in Table 3. In each trial, the time of the intervention is recorded and compared with what was expected. For example, if the time was between the realistic and the pessimistic time, the value "3" (from Table 4) is recorded in the column "Intervention rapidity". Similarly, a value of intervention quality is chosen from Table 5. The same action is used for the remaining seven trials.

From the achieved results, the coefficients of each factor of equation (1) can easily be calculated for the two chosen responses. Consequently, the model obtained from each response can be graphically represented so that we can come out with graphs to determine which competencies should be chosen for a given maintenance intervention fixing in advance the desired rapidity and quality of the intervention.

To illustrate these theoretical notions, we will present a practical study.

V. RESULTS AND DISCUSSION

The case study was conducted in an industrial enterprise. For confidential reasons, we cannot give details about this company. The experimentation must go through the following steps:

- Step 1: Determine the level of each actor based on the questionnaire in Annex 1
- Step 2: Determine the high and low level of each factor in the experimental design.
- Step 3: Choose eight different maintenance tasks that will be called "pilots" to start the experiment
- Step 4:Perform the experimentation;
- Step 5:Calculate the coefficients of each response and generate the final model;
- Step 6: Represent graphically the results.

The questionnaire was largely based on the ISO/IEC 17024 in the sense of properly allocating skills according to three levels of maintenance. This questionnaire was submitted to the head of the company and the results obtained are grouped in Table 6:

Factors	Low Level (-)	High Level (+)
Level 1 of maintenance	1	10
Level 2 of maintenance	0.3	7
Level 3 of maintenance	2	7

Table 6. Factors and field of study

Subsequently, we selected eight maintenance tasks to do the experiment. The results are summarized in Table 7:

Trials	**Factor 1**	**Factor 2**	**Factor 3**	**Intervention rapidityR_1**	**Intervention quality R_2**
1	**-1**	**-1**	**-1**	**4**	**3**
2	**+1**	**-1**	**-1**	**4**	**3**
3	**-1**	**+1**	**-1**	**3**	**3**
4	**+1**	**+1**	**-1**	**2**	**1**
5	**-1**	**-1**	**+1**	**3**	**2**
6	**+1**	**-1**	**+1**	**2**	**1**
7	**-1**	**+1**	**+1**	**1**	**1**
8	**+1**	**+1**	**+1**	**1**	**1**
Level -	**1**	**0.3**	**2**		
Level +	**10**	**7**	**7**		

Table 7. Experimentation matrix

The results of this experiment allowed us to determine the two following models corresponding to the rapidity and quality of intervention respectively:

$$R_1 = 2.5 - 0.5x_1 - 0.75x_2 - 0.75x_3 + 0.25x_1x_2x_3 \qquad (2)$$

$$R_2 = 1.875 - 0.375x_1 - 0.375x_2 - 0.625x_3 - 0.125x_1x_2 + 0.125x_1x_3 + 0.125x_2x_3 + 0.375x_1x_2x_3 \qquad (3)$$

To facilitate the representation of these two models in 2D, four graphs (Figs. 3, 4, 5 and 6) corresponding to the four rapidity responses 1, 2, 3 and 4 and three graphs (Figs. 7, 8 and 9) corresponding to the three responses of quality of intervention:

From the results of these graphs it is easy to deduce the skills to be chosen for a given intervention. For example, supposing we have arranged a maintenance intervention with a total margin other than zero. In other words, the intervention can be performed between the realistic and the pessimistic expected times. This corresponds to a rapidity of 3 (from Table 4). To do this, according to Fig. 4, it is possible to choose a maintenance actor having at least the following notes in the three maintenance levels:

- In Level 1: 10
- In Level 2: 2.53
- In Level 3: 5.62

Moreover, if we want quality intervention equals to 3, we must choose an actor with at least the following marks:

- In level 1: 2
- In level 2: 1.27
- In level 3: 2.71

If we have a critical maintenance task (total margin of the task = 0), we must choose an intervention rapidity of 1 because a delay on this task can cause losses. To do this, we must use figure 6 corresponding to the rapidity 1. Indeed, the actors that must be assigned to this task must possess, for example, at least the following marks:

- In level 1: 7
- In level 2: 0.5
- In level 3: 4

As a synthesis, it is clear that the established models are operational and can provide a very good help and support to the industrialists to affect its actors to the different daily tasks in an optimal way.

In addition, most of the research work that deals with the improvement of the quality of the maintenance process is concerned with the managerial and organizational aspects. However, the tool proposed in this paper deals with the operational part of the maintenance process.

Moreover, this tool is easily applicable for any manufacturing company because it does not require specific conditions and assumptions. Each company, from its maintenance activities and its human skills, can establish skills identification charts.

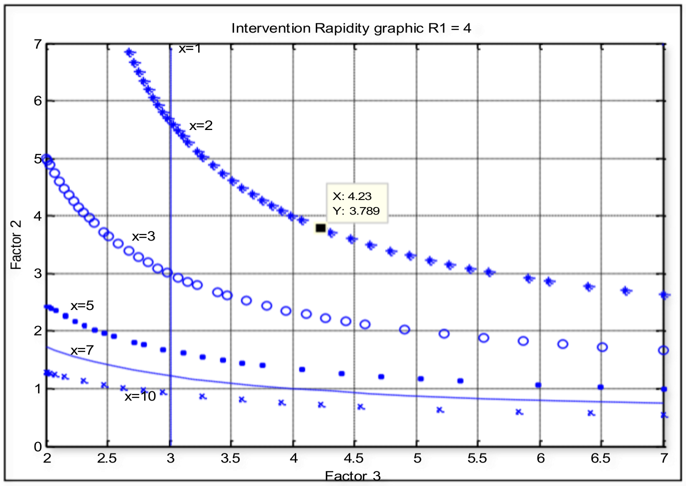


Fig. 3. Intervention rapidity graphic R1 = 4

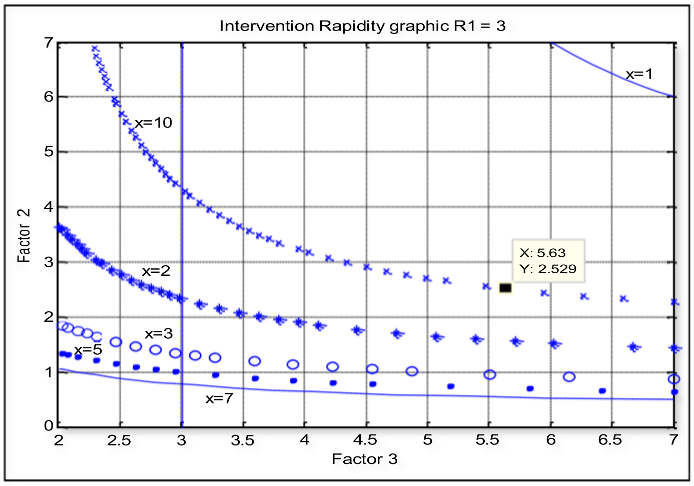


Fig. 4. Intervention rapidity graphic R1 = 3

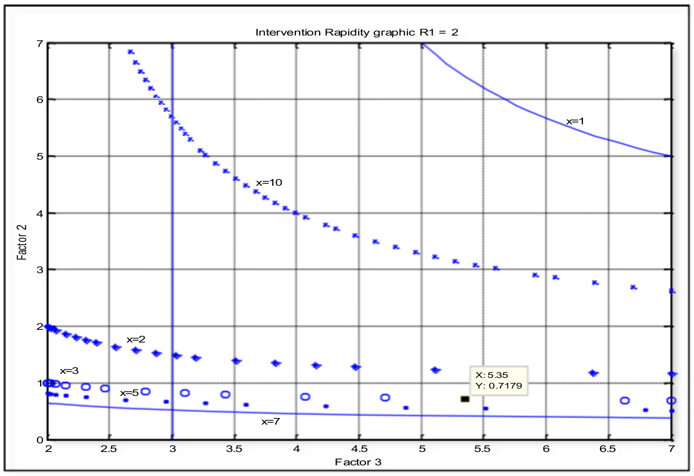


Fig. 5. Intervention rapidity graphic R1 = 2

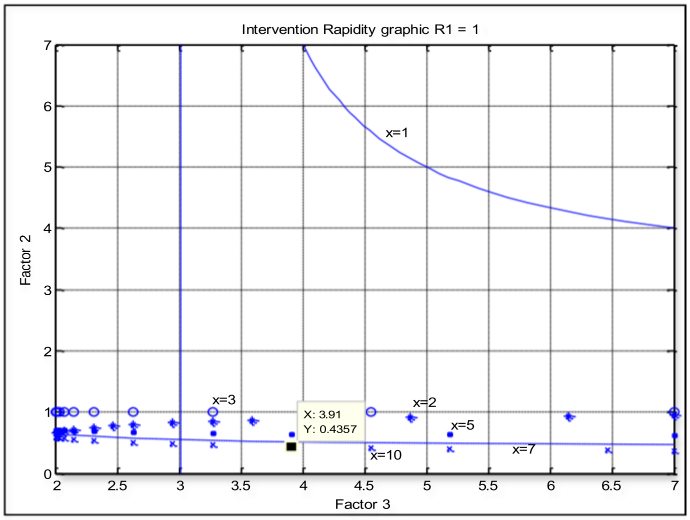


Fig. 6. Intervention rapidity graphic R1 = 1

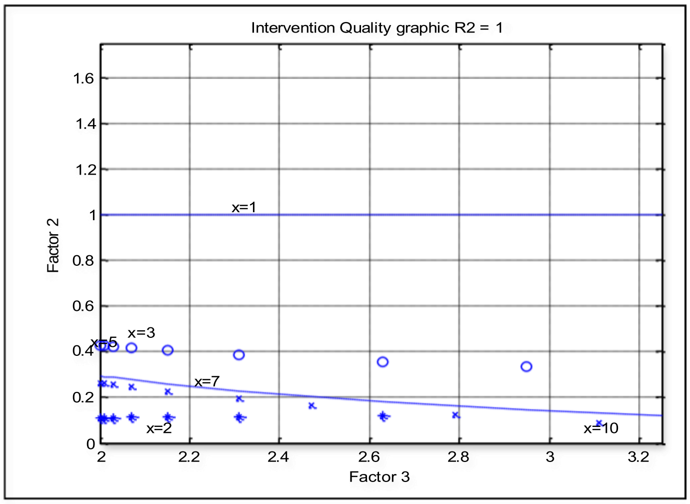


Fig. 7. Intervention quality graphic R2 = 1

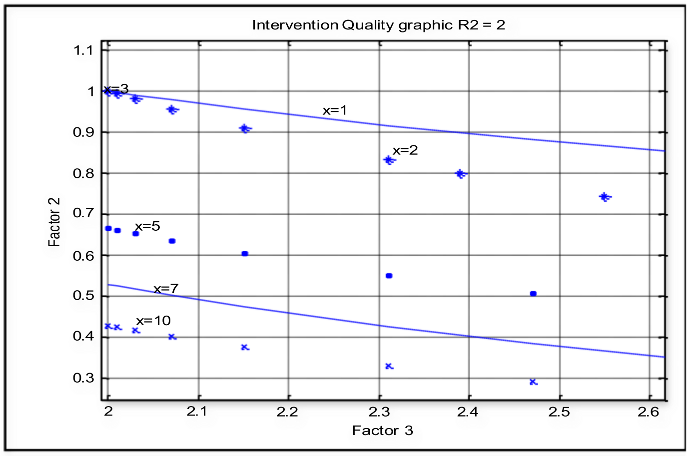


Fig. 8. Intervention quality graphic R2 = 2

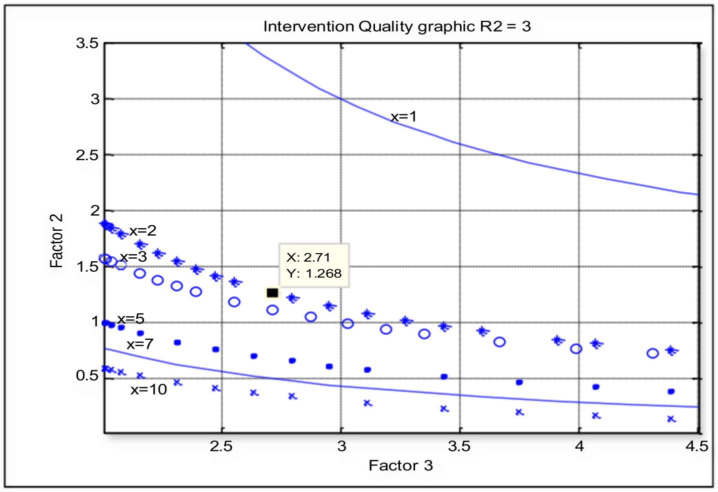


Fig. 9. Intervention quality graphic R2 = 3

VI. CONCLUSION

Taking into account the level of competence of the actors in the daily activities of the company is a basic factor and a fundamental necessity. It helps manufacturers to overcome the various difficulties and constraints linked to the tasks to be performed on one hand and to use and correctly allocate the actors to the different activities according to the level of competence and qualification of each one of them on the other hand. It is from this perspective that we are interested in suggesting a support tool in the choice of the skills to be assigned to the various maintenance tasks according to the rapidity of execution and the quality of intervention desired. Such a tool was based on the practice of the experimental design.

We began by drawing up the state of the art on the competence concept. We have presented, among other things, a few definitions of the concept and the main works carried out in this field in order to be able to situate the problem. Then, we presented the objectives and hypotheses of the study. This led us to set the objectives to be achieved being: the rapidity and the quality of maintenance intervention. Then, we sought the factors that can influence the responses as well as their levels. Subsequently, we chose the appropriate experimental design for such a study. To test the efficiency of the proposed tool we conducted a practical study.

Such a study led to the plot of the graphs corresponding to the responses chosen from the models obtained from the experiments carried out. These graphs help as a basic tool to select and appropriately assign actors to the various tasks and maintenance interventions.

In terms of prospects, we intend to extend this study to cover other fields and to conduct experiments in several industrial sectors in order to come out with specific charts for each sector serving as a reference for the choice and the allocation of the actors to the different activities.

REFERENCES

[1] Anne C Crook & Julian R Park, (2007). A Tool for Mapping Research Skills in Under graduate Curricula. *Journal Bioscience Education* 9(1), 1-12 DOI: 10.3108/beej.9.1

[2] Berio, G. Harzallah, M, 2005. Knowledge management for competence management, *Journal of Universal Knowledge Management*, 0(1), 21-28.

[3] Boucher, X. Bonjour E, Grabot, B. (2007). Formalisation and use of competencies for industrial performance optimisation: A survey. *Comp In Industry* 58(2) 98-117.

[4] Boumane A. (2007). *Développement d'une méthodologie de gestion de compétences*. Thèse de doctorat de l'université sidi Mohammed Ben Abdellah Fès

[5] Boumane A., Talbi A., Tahon C., Bouami D., (2006). Contribution à la modélisation des compétences. *MOSIM'06 - du 3 au 5 avril 2006* – Rabat-Maroc

[6] Bryan A. Norman , Wipawee Tharmmaphornphilas , Kim Lascola Needy , Bopaya Bidanda & Rona Colosimo Warner, (2002). Worker assignment in cellular manufacturing considering technical and human skills. *International Journal of Production Research,* 40(6), 1479-1492, http://dx.doi.org/10.1080/00207540110118082

[7] Dakkak B. Talbi A. (2014). *Amélioration de la performance industrielle par le déploiement du SMI : Application aux PME Marocaines.* Presse Académique Francophone

[8] Elizabeth B. Kuranchie-M., Kwesi Amponsah-T. (2016). Employee Motivation and Work Performance: A Comparative Study of Mining Companies in Ghana. *Journal of Industrial Engineering and Management JIEM*, 9(2): 255-309

[9] Guillaume E. Vanot and Michelle Sergent (2005). Experimental Design in Microbiology. *Microbial Processes and Products*. Springer p514

[10] Guillermo Campos Ciro, Frédéric Dugardin, Farouk Yalaoui & Russell Kelly, (2015). Open shop scheduling problem with a multi-skills resource constraint: a genetic algorithm and an ant colony optimisation approach. *International Journal of Production Research* 54(16), 4854-4881, DOI: 10.1080/00207543.2015.1126371

[11] Harzallah, M. Berio, G. Vernadat, F. (2006). Analysis and modelling of individual competencies: towards better management of human resources, *IEEE Transactions on Systems, Man and Cybernetics: Part A: Systems and Humans* 36 (1), 187 - 207

[12] Howard J. Seltman, (2018). *Experimental Design and Analysis*. Available in http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

[13] Imren Markes, (2006). A review of literature on employability skill needs in engineering. *European Journal Of Engineering Education*, 31(6), 637-650, http://dx.doi.org/10.1080/03043790600911704

[14] ISO/IEC 17024 (2012), Conformity assessment - *General requirements for bodies operating certification of persons*

[15] Jacques Goupy, Lee Creighton (2006). *Introduction aux plans d'expériences*. Edition Dunod, France.

[16] Janique S., Pierre M. (2003). Accroitre la performance par l'enrichissement des taches et l'autonomie - l'exemple des techniciens de fabrication.

[17] *Revue de Gestion des Ressources Humaines*, Eska, pp.45-56.

[18] Jorge Moutinho, Fernanda Oliveira (2015). Optimizing maintenance processes on customer site in a decentralized organization based on multi-site teams. *International Journal for Quality Research* 9(1) 141–150

[19] Jung Eun Hong (2016). Identifying Skill Requirements for GIS Positions: A Content Analysis of Job Advertisements. *Journal of Geography,* 115(4), 147-158 DOI: 10.1080/00221341.2015.1085588

[20] Le Boterf G., (2004). *Construire les compétences individuelles et collectives : les réponses à 90 questions.* 3éme édition, Editions d'Organisation.

[21] Lévy-Lebyer C., (1997). *La gestion des compétences.* France, Edition d'organisation.

[22] Lindgren, R., O. Henfridsson, and U. Schultze, (2004). Design Principles for Competence Management Systems: A Synthesis of an Action Research Study. *MIS Quarterly*, 28 (3), 435-472.

[23] Luis Mendes & José Machado, (2015). Employees' skills, manufacturing flexibility and performance: a structural equation modelling applied to the automotive industry. *International Journal Of Production Research,* 53(13), 4087–4101, http://dx.doi.org/10.1080/00207543.2014.993772

[24] Marly Monteiro de Carvalho & Roque Rabechini Junior, (2015). Impact of risk management on project performance: the importance of soft skills. *International Journal of Production Research* 53(2), 321-340, DOI: 10.1080/00207543.2014.919423

[25] Nassazi A. (2013). *Effects of training on employee performance*. Business Economics and Tourism

[26] Nilesh Pancholi, Mangal Bhatt (2017). Quality enhancement in maintenance planning through non-identical FMECA approaches. *International Journal for Quality Research* 11(3) 603–626

[27] Z. Arsovski, M. Pavlovic, S. Arsovski (2008). Improving the quality of maintenance processes using information technology. *International Journal for Quality Research*, 2(2) 101-114

Annex 1: Questionnaire for determining the levels of competence in each level of maintenance

	Questions	**Affirmation**			
		Yes	**Rather yes**	**Rather no**	**No**
Level 1	Carry out daily cleaning operations				
	Carry out daily grease operations				
	Carry out lubrication operations				
	Carry out checking and parameters adjustment operations				
	Carry out bolt tightening operations				
	Carry out state supervision rounds				
	Carry out statement value or usage unit				
	Perform manual manoeuvring of mechanical elements				
	Adjusting and replacing wear items or deteriorated parts on simple and accessible components				
	Use personal protection equipment				
Total					
Level 2	Check the parameters on equipment while operating, using integrated measuring means in the property				
	Make simple adjustments				
	Check the breaking equipment (sensors, circuit-breakers, fuses) and safety devices				
	Read a mechanical and electrical diagram of an equipment				
	Reading troubleshooting logic diagram for re-cycling the equipment				
	Identify components and sub-assembly of production equipment				
	Perform simple troubleshooting				
	Replace individual wear components or deteriorated by standard exchange with common tooling				
	Identify and apply equipment and intervention safety rules				
	Having notions about the practice of maintenance types				
Total					
Level 3	Check and adjust using external measuring instruments				
	Carrying out statement of technical parameters using measurements of individual measuring equipment				
	Conduct preventive maintenance visits on complex equipment				
	Diagnose common stops and dysfunctions				
	Finding the information needed to carry out the intervention				
	Performing technical repairs				
	Replacing organs and components by standard exchange of general technicality with specific tooling				
	Conduct and develop corrective and preventive maintenance				
	Exercise maintenance of equipment at risk				
	Contribute to the improvement of an industrial system in order to optimize the maintenance				
Total					

Sustainable last mile transportation and smart city: a literature review

Ibrahim Boukallal[1], Fatima El Khoukhi[2]

[1] Laboratory IA, FS Meknes, Moulay Ismail University of Meknes, Morocco
ibrahim.boukallal@edu.umi.ac.ma
[2] Laboratory IA, FLSH Meknes, Moulay Ismail University of Meknes, Morocco
f.elkhoukhi@flsh.umi.ac.ma

Abstract: The Smart City (SC) considers the city as a complex adaptive system of citizens, resources, and services. The characteristics of dynamic change and complexity are key issues for planning sustainable Last Mile (LM) transportation. Therefore, improving LM transport and significantly reducing these externalities are very important challenges for researchers. New technologies, transportation means, innovative techniques, and organizational strategies can be used to deal more efficiently with LM transport in SC. In this review, we present a synthesis of the problem of LM transport in SC, identifying current and emerging interests of researchers worldwide in this research area, and highlighting some open avenues. In this paper, we've examined various aspects related to the issue of LM transport within SC. Our aim has been to provide a more precise definition of SC, explore its characteristics and different generations, delve into the concept of SC, and ultimately, present the scientific research concerning LM transport in relation to these subjects.

Keywords: *Last Mile Transportation, Smart City, New technologies, Sustainability, Literature review.*

1 Introduction

The number of people residing in urban regions is constantly growing [1, 2, 3]. However, the exponential growth in transportation brought about by globalization and the internet market has made it feasible to create an open market. You may buy products anywhere in the globe, and the majority of them are delivered to cities. The issue of LM delivery is being tackled in a creative way with the development of information transport systems (ITS), information and communication technologies (ICT), and innovative transport vehicles [6, 10, 47]. Due to the increasing urban population, metropolitan areas require a significant quantity of resources, freight, and services, which leads to a number of issues for society [9, 14, 36]. The transportation sector is one of the main sources of external costs brought on by human activity. The scientific literature considers externalities such as air pollution, noise pollution, traffic congestion, accidents, climate change, and the degradation of transportation infrastructure. Traffic and the existence of transportation infrastructure are only two examples of the many local factors that affect the consequences of transportation-related air pollution emissions, which are very location-specific. The health impacts that a unit increase in air pollution concentration has on both humans and the environment are used to quantify emissions.

One of the most significant problems in transportation planning is climate change, which is mostly driven by greenhouse gas (GHG) emissions into the atmosphere. Additionally, the city's extensive automobile usage creates gridlock, which costs money and time and drastically reduces the efficacy of the transportation system. The transportation and parking of automobiles, especially those used for freight, also contributes to the degradation and expansion of the infrastructure. Nonetheless, social and economic considerations heavily depend on transportation-related activities. The movement of people (by private and public transportation) and the movement of goods are the two main divisions of the transportation business. Both places employ a variety of modes of transportation, including air, train, road, and ship.

The idea of SC has gained popularity in wealthy, industrialized nations. The main elements of a Smart Community are solid waste management, sanitation, affordable and intelligent housing, a sustainable power source, efficient public transit, urban mobility, and elements that support a clean environment for inhabitants [28, 35, 39]. Information technology, the Internet of Things (IoT), digitalization, artificial intelligence (AI), and machine learning (ML) techniques are also used to efficiently and successfully finish each component of a SC [2]. The most popular means of transportation for LM delivery is the road, while alternative modalities are frequently employed according on the features of the cities: For example, where there is a river, ships and barges are usually used. In urban locations, road freight transport is the main modality responsible for delivery-related externalities. This is why a lot of study has been done to try to improve the efficiency of delivery logistics while also reducing the externalities associated with it. LM delivery is a major problem because of the volume of traffic generated in the metropolitan area. In order to transport products to urban shops, shippers need a large number of carriers and logistical service providers, which makes it typically unorganized and fragmented. This results in many routes, low vehicle filling rates, large externalities, and high system costs. Reducing these externalities and offering citizens high-quality services are the main challenges to the implementation of LM services [35]. The development of new transportation systems and services through ICT is one of the most important issues that contribute to the new idea of the smart city. This will result in a sustainable and effective use of resources. Citizens are the main players in a smart city concept: ICT technologies support human activity in a variety of fields, such as economic growth, social and human connections, education, and environmental protection. For the city to operate successfully, sustainably, and competitively, these technologies must be used in conjunction with innovative and knowledge-based tactics [4, 5, 7, 32].

Based on a comparative analysis of 48 studies addressing these subjects, this report provides an overview of the literature on LM transport in SCs. The study's goal is to give academics and professionals in the field a useful tool for understanding how recent developments may affect the externalities brought on by urban freight transportation. The study is divided into five sections: Section 2 defines the concept of a smart city; Section 3 summarizes the primary studies on last-mile transportation. Finally, the conclusions are presented in Section 4.

2 Smart city

2.1 Introduction to smart city

Due to organizational, cultural, and technological shifts, the idea of SC has gained traction in recent years. In order to imagine sustainable, pleasant, and environmentally friendly mobility, the SC is a city that has optimized its functioning in a number of areas, including governance and the production or provision of urban services adapted to the population's traditional and modern demands [9]. A SC enhances the way various intelligent components are integrated to raise the standard of living for all of its citizens. When it comes to smart service delivery, citizen happiness is the main characteristic of smart cities. As seen in Figure 1, a smart city's design is made up of a number of components, such as smart energy, smart transportation, smart healthcare, smart homes, smart security, smart industry, smart community, and smart education system. Cities are intelligent and efficient because of these factors. Additionally, to enable all facets of a SC's functionality and operation smoothly and effectively, sophisticated information technologies and digitalization—such as AI, IoT, and ML—are used. With the rise of smart gadgets, the concept of linking several devices to a sophisticated network has gained a lot of traction. The development of conventional communications, which linked thousands of sentient things, gave rise to the Internet of Things.

FIG. 1 – A generic composition of SC architecture [21]

IoT has sparked a lot of attention from many stakeholders, which has resulted in the development of applications like as smart homes, smart warehouses, smart healthcare, and SCs [17, 18]. Due to global urbanization, SCs have drawn special attention in recent decades. Cities are now more efficient on many levels, including procurement, energy, healthcare, and transportation, thanks to the use of ICT in urban operations. The digital city, information city, and telicity—all of which are abstractions of the previous models—are among the various cities that have benefited from SCs [19]. The working mechanism of SCs is the same as that of IoT as SCs are the application domain of IoT [29].

To increase the efficiency of municipal operations and citizen services, the concept of a SC integrates ICT with a range of physical items connected to the IoT network. Generally speaking, a SC is an urban area that uses ICT and other IoT-based sensors to gather data. In order to improve the quality of service and efficiency of municipal operations, the data may then be used to manage resources, assets, and services in an efficient manner. Data is collected from residents, devices, and assets to analyze and manage traffic and transportation systems, utilities, power plants, water supplies, trash disposal, crime prevention, information systems, hospitals, schools, libraries, and other city services. After then, this data is examined and processed. In [37], a thorough definition of a SC is provided as follows:

Definition. A SC is a developed modern metropolis that makes use of ICTs and other technologies to enhance the standard of living, competitiveness, and operational effectiveness of urban services while guaranteeing the availability of resources in terms of social, economic, and environmental issues for both current and future generations ([37]).

2.2 Characteristics of smart city

The SC is composed of four major attributes, including sustainability, intelligence, urbanization, and quality of life [22, 26], as shown in Figure 2. Additionally, it is important to note that certain issues such as government, the economy, and infrastructure are closely linked to several of these attributes.

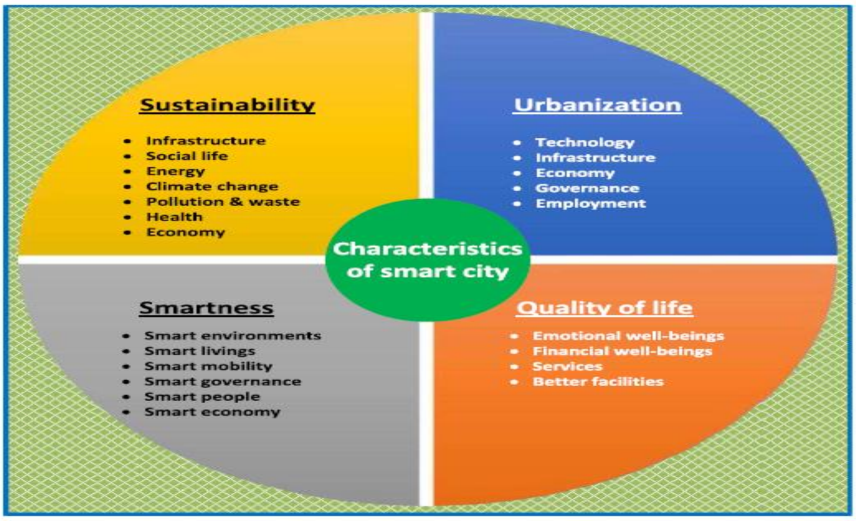


FIG. 2 – Characteristics of SC [19, 29]

Sustainability: A sustainable city (SC) addresses challenges related to waste management and pollution, energy and climate change, urban infrastructure and governance, social issues, economics, and health.

Urbanization: A SC urbanization may be measured in a many of ways, including through factors like infrastructure, technology, government, and the economy.

Quality of life: The mental and financial health of the populace can be used to gauge quality of life.

Intelligence: Enhancing the city's and its residents' social, economic, and environmental standards is the aim of SC intelligence. Some of the aspects of a city's intelligence that are frequently highlighted include smart surroundings, smart living, smart transit, smart government, smart people, and smart economics. Sustainable resource management, little pollution, and a beautiful, clean natural condition are all hallmarks of a smart ecosystem. Safety, health, standard of living, and access to cultural and educational resources are important markers of smart living. Smart mobility requires the availability of ICT infrastructure as well as safe, innovative, and sustainable transportation solutions. Smart governance is defined as transparent administration that involves people in decision-making and facilitates easy access to social and public services. Individuals who possess intelligence are imaginative, adept, receptive to new ideas, and willing to engage in society. Entrepreneurship, flexibility in the workplace, productivity, and international integration are factors that go into building a smart economy in a city.

2.3 Smart city generations

The adoption of technologies and advancements by cities is broken down into five stages in this section: technology (SC 1.0), city government, citizens, industry 4.0 (including 4G and 5G, electric vehicles, etc.), and artificial intelligence and cognitive computing (SC 5.0). The five stages of SC evolution are likewise depicted in Figure 3.

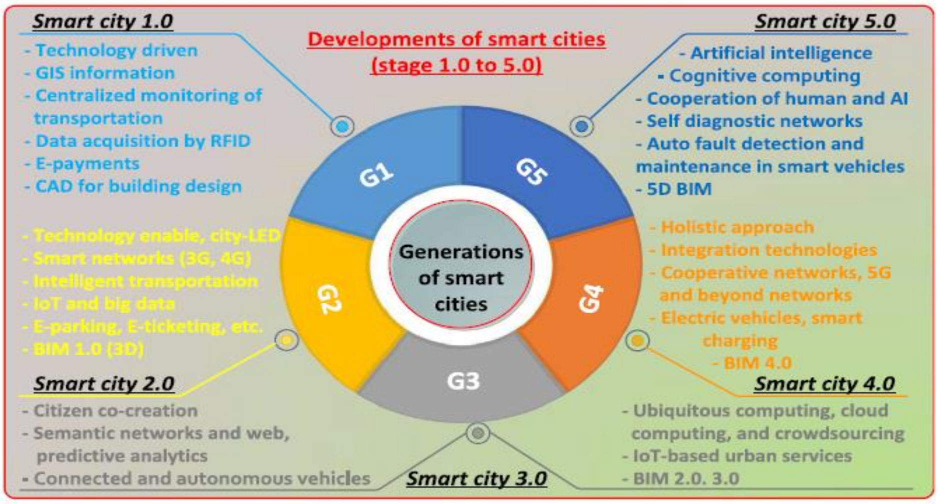


FIG. 3 – Five generations of SC [21]

Smart City 1.0: Technology companies are driving the adoption of their solutions in cities, which is a hallmark of SC 1.0. These cities are frequently criticized for their emphasis on technology and the significant influence of big businesses like Cisco and IBM. In contrast to the technology suppliers in SC 1.0, mayors and municipal councils, as well as a government agency, lead SC 2.0.

Smart City 2.0: When technical tools are specifically created to address issues like pollution, sanitation, health, and transportation in collaboration with citizens, SC 2.0 is applicable. Regretfully, only a small proportion of citizens find informal decision-making structures and assemblies appealing, and citizen involvement rates are low. While SC 3.0 is driven by citizen/user expectations, SC 1.0 and 2.0 are driven by government actions and technology, respectively.

Smart City 3.0: The government will develop government-specific user demands and serve as a facilitator for public opinion expression in SC 3.0. Therefore, a connected ecosystem (SC) encompasses all of the technologies, solutions, players, and audiences that are present in it. These include the Internet of Things (IoT), 5G connectivity, smart automotive and transportation, energy and utilities, public health and safety, artificial intelligence, and data analytics.

Smart City 4.0: The advantages of SCs are recognized to outweigh the costs of the city with the platformization of the city by embracing industrial revolution 4.0 [45]. The technological disruption of generation 1.0, the individualization of 2.0, and the engagement of 3.0 are all represented by SC 4.0, but it also adds two crucial success factors: a holistic approach and the difficulty of integrating solutions. The goal of the holistic approach is to combine newly created technologies with possibly undeveloped ones, as well as new and ancient technology. Furthermore, astute municipal leaders recognize that while SCs technologies can have a good impact on their communities, not every member of the community may experience these benefits equally. They also grasp the opportunities and limitations presented by new technology.

Smart City 5.0: The cooperative nature of SC 5.0 between humans and AI systems [33], allows it to seamlessly balance competing interests of various city stakeholders with all facets of existence. The method offered by City 5.0 can assist in reaching a "consensus" with various services and, more crucially, with the general public. Since it is a reflection of reality, it should take into account not just previous or present knowledge but also the dynamic interests, preferences, and limits of all parties involved in real time. These factors should be continually recognized, examined, turned into plans, put into action, and reviewed.

2.4 Literature review on smart city

The trend in the quantity of scientific papers on smart cities and associated topics, such sustainability and information technology, between 2020 and 2025 is depicted in Figure 4, which was collected from the Mendeley platform. The scientific community's increasing interest in these topics is demonstrated by this image, which shows a particularly noticeable trend between 2020 and 2023, when the number of publications increased dramatically. The quick development of new technologies, particularly AI and the IoT, which present previously unheard-of possibilities to revolutionize city administration, can be substantially attributed to this progression. Smarter, more effective administration of urban infrastructure, mobility, and public services is made possible by these technologies, which allow for real-time data collecting, advanced analysis, and more informed decision-making. The rise in publications on sustainability also indicates a rising interest in the social and environmental elements of smart cities. In fact, research is increasingly concentrating on solutions that improve social fairness and quality of life while lessening the ecological imprint of cities. This phenomenon is a reflection of a broader understanding of the global issues surrounding urbanization and the necessity of balancing sustainability and technological innovation in order to create the cities of the future. To put it briefly, the notable rise in scholarly works on smart cities and other subjects indicates both technological advancements in this area and a rising interest in addressing the social and environmental issues raised by global urbanization.

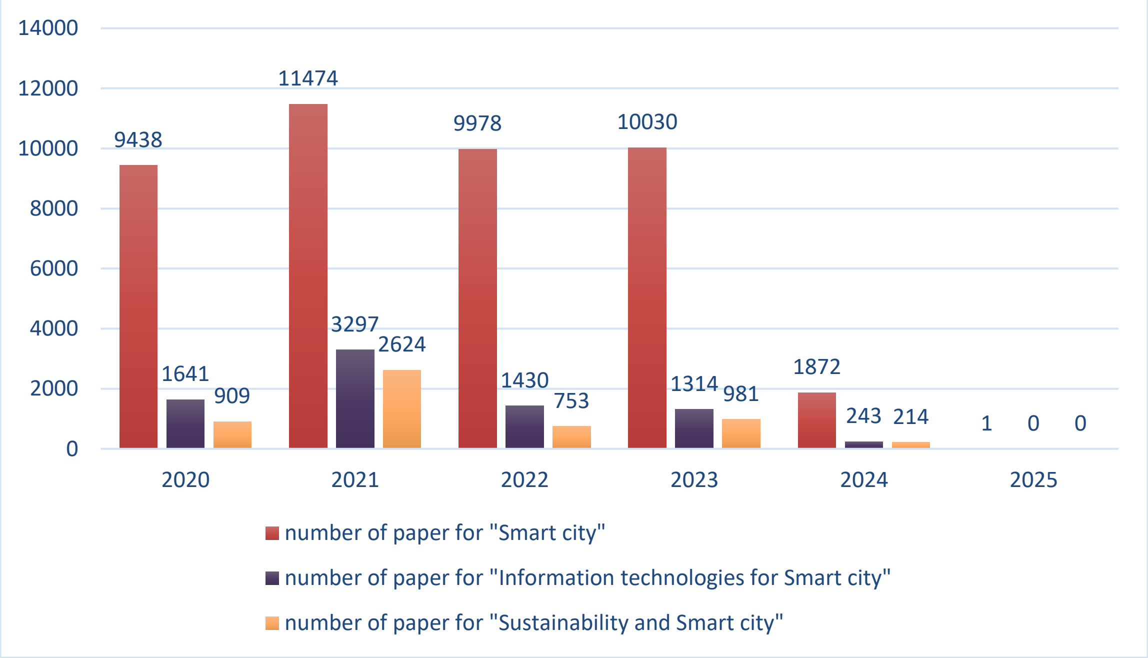


FIG. 4 – Search results for papers published in 2020-2025

A thorough overview of several studies on smart cities and the related technologies is given in Table 1. It offers a summary of current developments in this rapidly changing sector, making it a useful and useful tool for academics, decision-makers, and business experts. The table serves as a reference manual for the intricate realm of smart cities and associated technologies, making it a true information resource. Every study is succinctly described, including pertinent details like the authors' names to identify the top researchers in the area and the year of publication to find the newest trends and advances. The subjects discussed are also well-defined, offering a comprehensive grasp of the problems at hand, be it sustainability, urban infrastructure management, or new technology.

To help readers understand the methodological strategies and technology tools utilized in smart city initiatives, the table also lists the technologies utilized in each research. Lastly, readers who want to learn more about each topic or expand their knowledge may find complete references. In summary, this table provides a comprehensive overview of current technological advancements and applications in smart cities, in addition to summarizing the findings of several research studies. As such, it is an essential resource for anybody desiring to traverse this ever-changing cosmos.

Authors	Main topics	Technologies
Campisi T. et al. (2021) [10]	Smart Cities, Connected and Autonomous Vehicles (CAVs), Mobility Patterns	CAVs, Smart Infrastructure
Kauf S. (2019) [20]	Smart Logistics, Smart City Development	Logistics Optimization, Supply Chain Management
Ranieri L. et al. (2018) [32]	Last Mile Logistics, Externalities Cost Reduction, Sustainability	Logistics Optimization, Data Analytics
Xiao Z. et al. (2017) [44]	Online Retailing, Last Mile Delivery, Urban Logistics	E-commerce, Logistics Networks
Calvillo C. F. et al. (2016) [9]	Energy Management, Smart Cities, Planning	Renewable Energy Sources, Smart Grids
NATHALI B. et al. (2016) [29]	Smart Cities, Performance Improvement, Big Data Analytics	Big Data Analytics, Data Mining
Jin J. et al. (2014) [18]	Smart City, Internet of Things (IoT), Information Framework	IoT Sensors, Data Communication Networks

TAB. 1 – Main technologies used in smart city

In result, a review of current smart city research projects and scientific papers shows a dynamic of study that is becoming more diverse and intensive. Future cities are being shaped by the rise of new technologies like AI and the IoT, as well as a greater awareness of social and environmental challenges. For scholars, decision-makers, and professionals, the summary table provided here is a vital resource that gives them a better grasp of the most recent developments and the ways in which technology will affect future urbanization. In order to build cities that satisfy the requirements of their residents while upholding ecological and social imperatives, this research and innovation are crucial to the development of clever, sustainable, and inclusive solutions to the increasingly complex urban problems.

3 Last mile transport

The term "last mile" was initially used in the telecom industry to describe the final portion of a network. This final step in the delivery process is known as the final local warehouse, consolidation point, or distribution center in the context of the supply chain [44]. Synonyms like final-mile, home delivery, B2C distribution, and grocery delivery are also found in the literature [24]. LM transfer in SC is a relatively new and unexplored study subject in the literature. The views of many stakeholder groups, such as individuals, governments, shippers, receivers, transport companies, etc., are represented in a large number of publications [22].

Scholars claim that the supply chain's logistics management phase, which accounts for 28% of the total delivery cost, is the least efficient and environmentally friendly [1, 12, 15, 25]. According to a study of the literature on LM logistics in SCs and urban regions, the key advances to lower transportation costs and inefficiencies were collaborative urban logistics and improved transportation management and routing [32].

This transportation involves three distinct stakeholders: the supply side, which deals with the delivery of goods and is primarily represented by producers, online retailers, and courier, express, and parcel companies; the demand side, which deals with the demand for goods and is represented by individual consumers, businesses, and institutions; and the physical environment, which is within the jurisdiction of local government [3].

Planning for sustainable mobility is essential because of the LM's growing consumption and demand due to the ongoing growth of SCs. Therefore, it is essential in the SC to select different transportation providers for distribution services in a rational and reasonable way in order to boost efficiency and reduce expenses [40].

The fundamental concept underlying smart cities is the use of state-of-the-art information technology to achieve safe and intelligent city operation and administration and enhance people's quality of life. On the other hand, efficient distribution and transportation networks are among the most crucial components in improving the quality of service in smart cities [20].

Researching the mobility problem in smart cities is crucial because of this. Since 1959, the vehicle routing problem (VRP) has been the subject of much research [34, 38, 41, 42, 43]. Additionally, the expansion of e-commerce over the last two decades has increased the need for responsive omnichannel distribution to tackle the last mile challenge [26, 27]. Multimodality and flexibility in transportation networks will increase [31]. Several writers discuss the benefits of a simultaneous and integrated approach between client pickup and home delivery [23, 46].

3.1 Literature review on last mile transport

A thorough evaluation of the literature on vehicle uses in urban logistics, with an emphasis on sustainable practices, is shown in the Table 2 above. It gives a summary of current studies on important topics such last-mile delivery, route optimization, lowering greenhouse gas emissions, and using new technology to boost urban delivery efficiency. The key trends influencing the direction of urban logistics are shown in this table, which may be thought of as a roadmap of the most recent developments in this area. Since electric cars are thought to be a major factor in making delivery less polluting and more environmentally friendly, researchers are especially interested in them. In addition, the application of cutting-edge technology, such drones and driverless cars, is starting to show promise as a means of streamlining delivery procedures and lowering expenses and carbon emissions.

Other significant topics covered by the research in this table include route planning, cooperation amongst various industry participants (including customers, local government agencies, and transportation businesses), and the investigation of novel cooperative transportation models. The latter seek to reduce urban congestion and optimize resource utilization.

In summary, this figure illustrates how quickly the urban logistics industry is developing and how sustainability is becoming more and more significant in this industry. It attests to continuous attempts to improve delivery efficiency, cleanliness, and suitability for the modern urban difficulties. Nevertheless, there are still a lot of obstacles to overcome, especially when it comes to large-scale application, technology integration, and coordination. Thus, this table acts as a springboard for further studies meant to surmount these challenges and create an urban logistics system that is both inventive and really sustainable.

Authors	Main topics	Vehicles used
Fan et al. (2021) [41]	Time-dependent multi-depot green vehicle routing problem with time windows considering temporal-spatial distance	Electric Vehicles

Xiao et al. (2016) [38]	Heterogeneous green vehicle routing and scheduling problem with time-varying traffic congestion	Electric Vehicles and Alternative Fuel Vehicles
Mckinnon et al. (2016) [27]	Possible impact of 3D printing and drones on last-mile logistics: An exploratory study	Drones
Sassi et al. (2015) [42]	Vehicle Routing Problem with Mixed fleet of conventional and heterogenous electric vehicles and time dependent charging costs	Electric Vehicles
Erdoĝan et al. (2012) [15]	Green Vehicle Routing Problem	Electric Vehicles and Alternative Fuel Vehicles

TAB. 2 – Main vehicles used in LM transport

To sum up, the table provides a useful summary of recent developments and trends in urban logistics, emphasizing the creative and sustainable solutions that are revolutionizing the industry. A significant shift towards cleaner, more effective urban logistics that is better suited to the demands of contemporary cities may be seen in the emergence of electric cars, cutting-edge technology like drones, and cooperative approaches. Even if there has been a lot of development, there are still a lot of obstacles to overcome, particularly with regard to player coordination, technical integration, and ecological footprint reduction. This table is a crucial tool for directing future research and projects targeted at addressing these issues and assisting in the development of more intelligent and sustainable cities as it synthesizes the body of existing knowledge.

4 Conclusion and perspectives

This study provides a comprehensive overview of the many relevant issues raised by sustainability with regard to LM transport in SCs. It proposes a classification system for the topics and provides an overview of the main topics discussed in the literature. Our understanding of the numerous effects on the different players and their interactions is improved by this theoretical framework, which also makes it easier to identify the relevant concerns. LM logistics is a crucial phase in the supply chain. Researchers and practitioners are working together to improve delivery systems' efficacy and reduce associated externalities.

Urban transportation has a variety of externalities, such as air pollution from vehicle emissions, noise pollution from conventional engines and the sheer volume of traffic, accidents mainly brought on by distracted drivers, traffic jams from the city's high vehicle density, land use from vehicle sprawl, wear and tear on infrastructure from the weight and frequency of vehicles, and energy dependence from outdated power technologies and inefficient engines.

The needs of global trade and environmental standards must be met by LM transport, creating a dual challenge. According to this concept, innovation is essential to enabling the transition from the current logistics system configurations to more intelligent and sustainable ones. Innovations that can reduce externalities and affects can be grouped into five primary areas: new automobiles, proximity stations or points, collaborative and cooperative urban logistics, and better transportation management and routing. Each of them has advantages and disadvantages that can be regarded areas for more research.

For a SC to tackle the LM transport problem, innovative thinking will be required. One aspect of city management that necessitates a reevaluation of the movement of goods is the increase in express shipments, which is also a result of the growth of online sales, as well as new ICT and the Industry 4.0 paradigm, which allow the retrieval of massive amounts of data produced by devices, infrastructures, and automobiles. One of the main objectives of a smart city is to lessen the externalities brought on by last-mile delivery operations. All parties involved must be dedicated to achieving sustainable logistics in SCs, and the recommended solutions must be put into practice.

References

[1] Asghari, M., & Mirzapour Al-e-hashem, S. M. J. Green vehicle routing problem: A state-of-the-art review. *International Journal of Production Economics*, *231*(June 2020), 107899. https://doi.org/10.1016/j.ijpe.2020.107899. 2021.

[2] Aslam, S., Herodotou, H., Mohsin, S. M., Javaid, N., Ashraf, N., & Aslam, S. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renewable and Sustainable Energy Reviews*, *144*(March). https://doi.org/10.1016/j.rser.2021.110992. 2021.

[3] Bandeira, R. A. M., D'Agosto, M. A., Ribeiro, S. K., Bandeira, A. P. F., & Goes, G. V. A fuzzy multi-criteria model for evaluating sustainable urban freight transportation operations. *Journal of Cleaner Production*, *184*, 727–739. https://doi.org/10.1016/j.jclepro.2018.02.234. 2018.

[4] Boukallal, I., Bouti, N., & El Khoukhi, F. *Le problème AF-VRP : une revue de la littérature*. La Conférence Internationale Sur Le Monitoring Des Systèmes Industriels (CIMSI 2022), 14-15 Décembre à Tanger, Maroc, 1–5. 2022.

[5] Bouti, N., Boukallal, I., & El Khoukhi, F. *Optimization of Inventory Management: A Literature Review.* Chapter and Conference Paper, Digital Technologies and Applications Proceedings of ICDTA'13 (Fez, Morocco), Lecture Notes in Networks and Systems, 669 LNNS, 923–933. https://doi.org/10.1007/978-3-031-29860-8_92. 2023.

[6] Bouti, N., Boukallal, I., & El Khoukhi, F. *Sustainable Urban Logistics and Supply Chain 4.0: A Literature Overview and Future Directions.* Chapter and Conference Paper, Proceeding of GOL'24 "Smart Sustainable and Green Logistics", Volume 1, Part of the Lecture Notes in Networks and Systems Book Series (LNNS 1104), 335–344. https://doi.org/10.1007/978-3-031-68628-3_32. 2024.

[7] Bouti, N., & El Khoukhi, F. *Smart Supply Chain Management: A Literature Review*. Chapter and Conference Paper, Digital Technologies and Applications Proceedings of ICDTA'13 (Fez, Morocco), Lecture Notes in Networks and Systems Book Series, LNNS 668(1), 896–904. https://doi.org/https://doi.org/10.1007/978-3-031-29857-8_89. 2023.

[8] Brenner, N., & Schmid, C. The "urban age" in question. *International Journal of Urban and Regional Research*, *38*(3), 731–755. https://doi.org/10.1111/1468-2427.12115. 2014.

[9] Calvillo, C. F., Sánchez-Miralles, A., & Villar, J. Energy management and planning in smart cities. *Renewable and Sustainable Energy Reviews*, *55*, 273–287. https://doi.org/10.1016/j.rser.2015.10.133. 2016.

[10] Campisi, T., Severino, A., Al-rashid, M. A., & Pau, G. The development of the smart cities in the connected and autonomous vehicles (Cavs) era: From mobility patterns to scaling in cities. *Infrastructures*, *6*(7). https://doi.org/10.3390/infrastructures6070100. 2021.

[11] Cohen, B. Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society*, *28*(1–2), 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005. 2006.

[12] Erdoĝan, S., & Miller-Hooks, E. A Green Vehicle Routing Problem. *Transportation Research Part E: Logistics and Transportation Review*, *48*(1), 100–114. https://doi.org/10.1016/j.tre.2011.08.001. 2012.

[13] Fan, H., Zhang, Y., Tian, P., Lv, Y., & Fan, H. Time-dependent multi-depot green vehicle routing problem with time windows considering temporal-spatial distance. *Computers and Operations Research*, *129*, 105211. https://doi.org/10.1016/j.cor.2021.105211. 2021.

[14] Freudendal-Pedersen, M., Kesselring, S., & Servou, E. What is smart for the future city? Mobilities and automation. *Sustainability (Switzerland)*, *11*(1). https://doi.org/10.3390/su11010221. 2019.

[15] Gevaers, R., Van De Voorde, E., & Vanelslander, T. Characteristics and typology of lastmile logistics from an innovation perspective in an Urban context. *City Distribution and Urban Freight Transport: Multiple Perspectives*, *January*, 56–71. https://doi.org/10.4337/9780857932754.00009. 2011.

[16] Golinska-Dawson, P., & Sethanan, K. Sustainable Urban Freight for Energy-Efficient Smart Cities—Systematic Literature Review. *Energies*, *16*(6). https://doi.org/10.3390/en16062617. 2023.

[17] Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K. S. The internet of things for health care: A comprehensive survey. *IEEE Access*, *3*, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951. 2015.

[18] Jin, J., Gubbi, J., Marusic, S., & Palaniswami, M. An information framework for creating a smart city through internet of things. *IEEE Internet of Things Journal*, *1*(2), 112–121. https://doi.org/10.1109/JIOT.2013.2296516. 2014.

[19] Joshi, S., Mohanty, S. P., & Kougianos, E. Everything You Wanted to Know about PUFs. *IEEE Potentials*, *36*(6), 38–46. https://doi.org/10.1109/MPOT.2015.2490261. 2017.

[20] Kauf, S. Smart logistics as a basis for the development of the smart city. *Transportation Research Procedia*, *39*(2018), 143–149. https://doi.org/10.1016/j.trpro.2019.06.016. 2019.

[21] Khan, A., Aslam, S., Aurangzeb, K., Alhussein, M., & Javaid, N. Multiscale modeling in smart cities: A survey on applications, current trends, and challenges. *Sustainable Cities and Society*, *78*(March 2021), 103517. https://doi.org/10.1016/j.scs.2021.103517. 2022.

[22] Kiba-Janiak, M., Marcinkowski, J., Jagoda, A., & Skowrońska, A. Sustainable last mile delivery on e-commerce market in cities from the perspective of various stakeholders. Literature review. *Sustainable Cities and Society*, *71*(April). https://doi.org/10.1016/j.scs.2021.102984. 2021.

[23] License, N. C. C. B. I., Zhou, L., Vigo, D., & Wang, X. A Multi-Depot Two-Echelon Vehicle Routing Problem with Delivery Options Arising in the Last Mile Distribution, *European Journal of Operational*. *265*(2), 765–778. 2018.

[24] Lim, S. F. W. T., Jin, X., & Srai, J. S. Consumer-driven e-commerce: A literature review, design framework, and research agenda on last-mile logistics models. *International Journal of Physical Distribution and Logistics Management*, *48*(3), 308–332. https://doi.org/10.1108/IJPDLM-02-2017-0081. 2018.

[25] Lin, C., Choy, K. L., Ho, G. T. S., Chung, S. H., & Lam, H. Y. Survey of Green Vehicle Routing Problem: Past and future trends. *Expert Systems with Applications*, *41*(4 PART 1), 1118–1138. https://doi.org/10.1016/j.eswa.2013.07.107. 2014.

[26] Mckinnon, A. C. The possible impact of 3D printing and drones on last-mile logistics: An exploratory study. *Built Environment*, *42*(4), 617–629. https://doi.org/10.2148/benv.42.4.617. 2016.

[27] Melacini, M., Perotti, S., Rasini, M., & Tappia, E. E-fulfilment and distribution in omni-channel retailing: a systematic literature review. *International Journal of Physical Distribution and Logistics Management*, *48*(4), 391–414. https://doi.org/10.1108/IJPDLM-02-2017-0101. 2018.

[28] Nam, T., & Pardo, T. A. Conceptualizing smart city with dimensions of technology, people, and institutions. *ACM International Conference Proceeding Series*, *June*, 282–291. https://doi.org/10.1145/2037556.2037602. 2011.

[29] NATHALI, B., JUNG, C., KANG, J., SEO, J., KIM, J., HAN, K., KHAN, M., JIN, S., & YOON, Y. *Planning of smart cities Performance improvement using big data analytics approach*. *December*, 51–55. https://doi.org/10.15224/978-1-63248-113-9-11. 2016.

[30] Ouba, A., & Khoukhi, F. El. *Supply chain risk management: A literature review*. International Conference on Services and Industry of the Future (ICSIF2023), October, 24-26, 2023, Rabat, Morocco. 2023.

[31] Prause, G., Atari, S., Prause, G., & Atari, S. On sustainable production networks for Industry 4.0 To cite this version : HAL Id : hal-01860909. *Entrepreneurship and Sustainability Issues, Entrepreneurship and Sustainability Center*, *4*(4), 421–431. 2017.

[32] Ranieri, L., Digiesi, S., Silvestri, B., & Roccotelli, M. A review of last mile logistics innovations in an externalities cost reduction vision. *Sustainability (Switzerland)*, *10*(3), 1–18. https://doi.org/10.3390/su10030782. 2018.

[33] Rosemann, M., Becker, J., & Chasin, F. City 5.0. *Business and Information Systems Engineering*, *63*(1), 71–77. https://doi.org/10.1007/s12599-020-00674-9. 2021.

[34] Sassi, O., Cherif, W. R., & Oulamara, A. Vehicle Routing Problem with Mixed fleet of conventional and heterogenous electric vehicles and time dependent charging costs. *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, *9*(3), 167–177. 2015.

[35] Shee, H. K., Miah, S. J., & De Vass, T. Impact of smart logistics on smart city sustainable performance: an empirical investigation. *International Journal of Logistics Management*, *32*(3), 821–845. https://doi.org/10.1108/IJLM-07-2020-0282. 2021.

[36] Shen, S., Wei, Z. Q., Sun, L. J., Su, Y. Q., Wang, R. C., & Jiang, H. M. The shared bicycle and its network—internet of shared bicycle (IoSB): A review and survey. *Sensors (Switzerland)*, *18*(8), 1–24. https://doi.org/10.3390/s18082581. 2018.

[37] Silva, B. N., Khan, M., & Han, K. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society*, *38*(August 2017), 697–713. https://doi.org/10.1016/j.scs.2018.01.053. 2018.

[38] Subramanian, A., Penna, P. H. V., Uchoa, E., & Ochi, L. S. A hybrid algorithm for the Heterogeneous Fleet Vehicle Routing Problem. *European Journal of Operational Research*, *221*(2), 285–295. https://doi.org/10.1016/j.ejor.2012.03.016. 2012.

[39] Vinod Kumar, T. M., & Dahiya, B. Smart Economy in Smart Cities (Issue November). https://doi.org/10.1007/978-981-10-1610-3_1. 2017.

[40] Wang, H., Li, W., Li, M., Yang, X., Wang, Z., Zhao, Z., & Wang, L. Intelligent selection of delivery parties for fresh agricultural product based on third-party logistics in smart city. *Sustainable Energy Technologies and Assessments*, *52*(PB), 102151. https://doi.org/10.1016/j.seta.2022.102151. 2022.

[41] Wang, H., Li, W., Zhao, Z., Wang, Z., Li, M., & Li, D. Intelligent Distribution of Fresh Agricultural Products in Smart City. *IEEE Transactions on Industrial Informatics*, *18*(2), 1220–1230. https://doi.org/10.1109/TII.2021.3073128. 2022.

[42] Wang, Y., Li, Q., Guan, X., Xu, M., Liu, Y., & Wang, H. Two-echelon collaborative multi-depot multi-period vehicle routing problem. *Expert Systems with Applications*, *167*(August 2020), 114201. https://doi.org/10.1016/j.eswa.2020.114201. 2021.

[43] Xiao, Y., & Konak, A. The heterogeneous green vehicle routing and scheduling problem with time-varying traffic congestion. *Transportation Research Part E: Logistics and Transportation Review*, *88*, 146–166. https://doi.org/10.1016/j.tre.2016.01.011. 2016.

[44] Xiao, Z., Wang, J. J., Lenzer, J., & Sun, Y. Understanding the diversity of final delivery solutions for online retailing: A case of Shenzhen, China. *Transportation Research Procedia*, *25*, 985–998. https://doi.org/10.1016/j.trpro.2017.05.473. 2017.

[45] Yun, Y., & Lee, M. Smart City 4.0 from the perspective of open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(4). https://doi.org/10.3390/joitmc5040092. 2019.

[46] Zhou, L., Wang, X., Ni, L., & Lin, Y. Location-routing problem with simultaneous home delivery and customer's pickup for city distribution of online shopping purchases. *Sustainability (Switzerland)*, *8*(8). https://doi.org/10.3390/su8080828. 2016.

[47] Zygiaris, S. Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems. *Journal of the Knowledge Economy*, *4*(2), 217–231. https://doi.org/10.1007/s13132-012-0089-4. 2013.

The supply chain management during the COVID-19 pandemic: A review on Mathematical models

Amine Ouba[1], Fatima El Khoukhi[2]

[1] Laboratory IA, FS Meknes, Moulay Ismail University of Meknes, Morocco

am.ouba@edu.umi.ac.ma

[2] Laboratory IA, FS Meknes, Moulay Ismail University of Meknes, Morocco

f.elkhoukhi@flsh.umi.ac.ma

Abstract

In Supply Chain Management (SCM), decision-makers use various methods to achieve their goals, including minimizing costs related to transportation, production, and other supply chain stages. Sustainability is a key focus, with efforts to reduce CO2 emissions from production and transportation and enhance employment opportunities. Improving customer service levels is another priority, achieved by reducing the distance between consumers and service centers to enable faster deliveries, lower prices, and prevent product shortages. The COVID-19 pandemic exposed significant vulnerabilities in supply chains, particularly their susceptibility to disruptions. As a result, safeguarding supply chains against such risks has become a critical priority for decision-makers.

Among the quantitative approaches, mathematical modeling has proven particularly effective in creating resilient supply chains that can withstand risks, as demonstrated during the COVID-19 pandemic. This paper examines how mathematical modeling has been applied to address pandemic-related risks in SCM, drawing on two review papers with similar research focuses and analyzing nine papers that propose mathematical models for managing supply chains during the pandemic. It explores the application of multi-objective modeling to balance cost reduction, sustainability, and risk mitigation, focusing on constraints such as matching supply with demand, managing storage capacity, and minimizing transportation emissions.

Overall, this paper offers a comprehensive view of SCM during the pandemic, illustrating how mathematical modeling has effectively addressed both traditional goals and new challenges. It also discusses the long-term implications of developing resilient supply chains capable of handling probabilistic risks, to avoid the risks that they may come in the future.

Keywords: Supply chain risk management (SCRM), COVID-19, Optimization, Mathematical Modeling, Literature review.

1 Introduction

A supply chain (SC) is: “a network used to deliver products and services from raw material suppliers to end customers through an integrated flow of information, materials and finance” [33]. In other words, there's the notion of SCM, in which there are a number of objectives to be achieved:

- Ensuring end-customer satisfaction by minimizing lost sales, delivering products on time, and meeting quality standards, and complying with environmental conditions.

- Minimizing costs, including those associated with transportation, production, and fixed installations.

- Establishing a sustainable chain involves respecting environmental conditions, mitigating uncertain risks, ensuring the chain's resilience to potential disruptions, and verifying adherence to social standards to enhance employment opportunities [36].

Sustainability, is one of the new concepts developed in SC risk management, according to (Ivanov et al., 2023) [21] the pandemic motivated an extension of SC resilience (recovering from a disruptive event) to a more complex and longer-term notion we call SC sustainability. SC viability encompasses research at the intersection of resilience, adaptability and sustainability (Ivanov and Dolgui.,2020) [20]. Since 2020, the risk of a COVID-19 pandemic has succeeded in causing major disruptions (a sharp increase in demand, the closure of locations supplying raw materials resulting in a major shortage in the global market, a shortage of means of transport, a reduction in production capacity...[5]) to SC, which can result in the entire chain coming to a standstill. “The recent outbreak of COVID-19 has caused great uncertainty for SC networks worldwide” [21], (Schleifenheimer and Ivanov., 2024) [31] submitted an example of the risk caused by COVID-19 also extended to transportation, which was disrupted on several levels. Many countries imposed substantial restrictions on imports and exports, and border security was increased to ensure the adherence to limitations of in-country and cross-border movements. Disruptions also adversely affected the supply side, where shortages were rising for excipients and active pharmaceutical ingredients (APIs). The main impact resulted from export restrictions and bans imposed by China and India, two major producers of APIs and generics, and unexpectedly limited supply, as well as labor shortages and temporary closures. The demand increases and simultaneous hoarding and non-release of pharmaceutical supplies to buying countries multiplied the effect of the disruptions and led to an unforeseen erratic product supply.

During the COVID-19 pandemic, SCM managers were motivated to develop multi-objective and mono-objective mathematical models. These models aim to determine more accurate approaches for Sc to become proactive, rather than reactive. This paper will provide a global synthesis of important papers existing in the literature that propose methodologies for risk management especially those addressing the impacts of the COVID-19 pandemic. Our focus here is on quantitative mathematical methodologies, particularly multi-objective or mono-objective optimization, to create a viable SC and achieve its main objectives.

As shown in Figure 1, most of the research existing in the work of (Montoya-Torres et al., 2023)[26], used qualitative approaches for managing the SC during COVID-19, being the most representative the conceptual or theoretical papers (51%), followed by case studies (32%), and the analysis of policies (15%), and surveys or interviews (13%). On the other hand, quantitative approaches are used by 44.3% of the short-listed papers. Mathematical models (16%) and analytical techniques (including other mathematical/analytical formulations (17%), simulation models (6%), algorithmic approaches (3%), and machine learning (2%)) are the most important ones, we have 33% the quantitative approaches are the modeling mathematics for SCM during COVID-19, one of the tests existing in the literature which shows the importance of mathematical modeling during COVID-19.

The main objective of our study is to analyze and demonstrate how mathematical modeling has been employed to create resilient and robust supply chains during the COVID-19 pandemic. It aims to highlight the effectiveness of these models in addressing key challenges such as cost reduction, sustainability, and risk mitigation while ensuring supply chain continuity. Additionally, the study seeks to provide insights into the potential of mathematical approaches for enhancing supply chain resilience against future disruptions.

The rest of our paper is structured as follows: The first section presents an introductory overview of key concepts in SCM, including the challenges posed by COVID-19. A literature review of a set of papers utilizing single-objective or multi-objective mathematical modeling to address SCM challenges during the pandemic, is given in section II, the third section corresponds to presents the Impacts of COVID-19 on the supply chain, to determine the role of modeling mathematics in SCM during COVID-19 we constructed a fourth section and finally a conclusion presenting our findings and future directions.

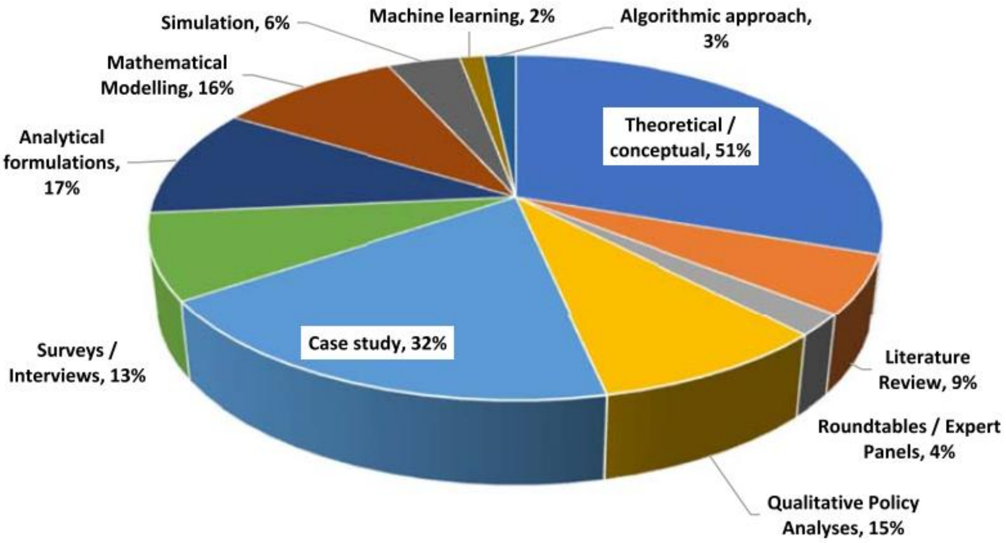


FIG. 1 – Most frequently used research methodology in the short-listed papers [26]

2 Literature review

Mathematical modeling has demonstrated to the world that it is one of the most important management tools in the field of supply chains. It has successfully addressed several management challenges, such as cost reduction, sustainability, and risk minimization. During the COVID-19 period, multi-objective optimization became highly competitive in the literature for optimizing the increased risks during this time.

This section delves into summaries of a set of existing papers in the literature answering the question of SCM during the COVID-19 epidemic using quantitative multi-objective or mono-objective mathematical modeling methodology.

(Liu et al., 2023) [24] proposed a stochastic mathematical model with two objectives: minimization of the risk disturbance on the manufacturing plant and minimizing the costs of the actions chosen by the suppliers. In this paper, the authors developed the Bayesian network method to address this transformation of risk disturbance between SC members which we call the ripple effect.

The COVID-19 period had social impacts on humanity, such as the loss of job opportunities, To address this challenge (Tirkolaee et al., 2023) [36] proposed to solve this problem by a deterministic mathematical model with two objectives minimizing total costs and maximizing job opportunities in SC, in a chain that transforms blood from a donation center to local hospitals.

Vaccination Sc in the COVID-19 period are addressed by (Tang et al., 2022) [33] Their paper studies the multi-period vaccination planning problem, which simultaneously optimizes two objectives: (i) minimizing the total operational cost, including the fixed cost of opening the vaccination site, the capacity selection cost, the replenishment cost and the inventory cost; (ii) minimizing the total population travel distance (to improve beneficiary comfort, i.e. service level). The decisions to be made, for each period, are 1) which vaccination sites to open; 2) how many vaccination stations to set up in each open site; 3) how much resupply in each site; and 4) how to allocate the population to the open sites.

An integrated modeling approach is proposed in the study (Azani et al., 2021) [5] which addresses the food SC consisting of production, distribution and consumption activities in a more comprehensive way about the COVID-19 epidemic. A mixed integer nonlinear programming (MINLP) mathematical planning model is proposed to reduce costs and environmental impact. Two scenarios are considered, the areas covered by the food SC are divided into high-risk and low-risk zones of propagation of the pandemic.

Given these limitations in the existing literature, the study by (Paul et al., 2022) [30] aims to develop a quantitative recovery model to manage the impacts of the multidimensional uncertainty of the COVID-19 pandemic. More specifically, the authors propose the following research objectives: (i). Develop a mathematical model for a three-stage SC to recover from the multidimensional impacts of the COVID-19 pandemic, such as increased demand, reduced production capacity, and reduced supply capacity; (ii). Consider simultaneously the uncertainties of demand, supply, and production capacity; and (iii). Develop an efficient approach to solving the model for both small- and large-scale problems.

In the context of public health, drug SC networks are faced with severe shortage problems in many situations, such as during the COVID-19 period. Drug shortages can occur due to manufacturing problems, lack of infrastructure and immediate reaction mechanisms, (Lozano-Diez et al., 2020) [25] proposing a single-objective model with constraints to solve this drug shortage considering the disruptions caused by various risks. The case study presented in the paper focuses on the COVID-19 pandemic's impact on Mexico. In the literature, we have a lot of important papers presenting the quantitative method of mathematics modeling multi or mono-objective for SCM during COVID-19 notably [37,38,27], and [16]. Table 1 presents the characteristics of the papers using mathematical modeling to manage chains during COVID-19 that we have taken to answer our research question.

TAB. 1 – The synthesis of some work exists in the literature

Author	Model type	Determinist	Stochastic	Objectives
(Liu et al., 2023) [24]	MDB DBN MIP		●	Minimizing the manufacturer's disruption risk and the expected total action cost
(Tirkolaee et al., 2023) [36]	MILP		●	Minimizing network costs and maximizing job opportunities
(Tang et al., 2022) [34]	MILP	●		Reducing the operational cost and the total travel distance.
(Azani et al., 2021) [5]	MINLP	●		Minimizing the total environmental impact Minimizing the total cost
(Paul et al., 2022) [30]	MILP		●	Maximizing the total profit
(Lozano-Diez et al., 2020) [25]	MILP	●		Minimizing cost
(Ala et al., 2024)[2]	MILP	●		Minimizing the cost Minimizing the total emission of CO2 Maximizing the social factors by creating jobs
(Shayannia et al., 2023)[32]	MILP	●		Economic objective function The social or environmental objective function Supply risk function Political stability objective function
(Fatemi et al., 2022)[15]	MINLP	●		Minimizing total costs Minimizing unfulfilled demands Reduce the waiting time at the factory Entrance

MDB: Markov decision process. MIP: Mixed-integer programming model.

DBN: Dynamic Bayesian network. MILP: Mixed-Integer Linear Programming.

MINLP: Mixed-Integer Non-Linear Programming.

In supply chain modeling, mathematical constraints are essential to ensure that the solutions generated by the model are feasible and align with real-world limitations. Table 2 presents some common categories of constraints in the studies cited:

TAB. 2 – Some common categories of constraints in modeling mathematical

Constraint area	Specific Constraint
Capacity	Storage Capacity: Limits on the amount of inventory a warehouse can hold. Transportation Capacity: Restricts the volume or weight that vehicles can carry.
Demand satisfaction	Ensures that customer demands are met.
Supply	Limits the maximum quantity available from suppliers.
Flow balance	Maintains balance between incoming and outgoing goods at a location.
Financial	Limits the total cost to remain within budget.
Time	Restricts delivery times to meet deadlines or lead times.
Binary/Integer	For decisions like whether to open a facility or select a route.
Service Level	Ensures the model meets customer service requirements.
Sustainability	Limits on emissions or waste.
Regulatory	Adherence to legal or contractual requirements, such as tariffs, quotas, or safety standards.

3 Types of problems found in the study by Chowdhury

The SC received a lot of problems during the outbreak of the pandemic COVID-19, we have in the literature the works that can class these problems, (Chowdhury et al., 2021)[12] In his work presented the table illustrates all classes of the impacts of COVID-19 on the supply chain.

TAB. 3 – The class on the impacts of COVID-19 on the SC[12]

Impacted area	Specific impact
Demand	Demand spikes for essential products. Shortage of essential products. Loss of security concerning essential items. Failure of on-time delivery. Declining demand for non-essential products. Ambiguity or difficulty in forecasting.
Supply	Shortage of material supply/supply-side shock/supply disruption.
Production	Production disruption and backlog. Reduced production capacity. Unavailability of workforce. Obsolescence and impairment of machinery and capital assets.
Transportation and logistics	Delays in transportation and distribution. Lack of international transportation/trade. Loss/lack of physical distribution channels. Shift of distribution and logistics pattern (offline to online or blended).
Relationship	Reduced social interaction. Information ambiguity. Lack of supplier engagement/opportunistic behavior.

Impacts on internal, upstream, and downstream operations	The ripple effect on all the operations involved in supply chains. Supply chain collapse. Closure of facilities, including both companies' production facilities and the facilities of supply chain partners such as suppliers and distributors
Financial	Reduced supply chain financial performance (e.g., loss/ reduction of financial stability) Reduced cash inflow
Sustainability	Lack of focus on social and environmental sustainability practices/disruption of sustainability initiatives. Threats to the health and safety of the workforce. Contraction of the development of green and low-carbon energy sources Increase in waste. Increased in recyclable materials.

On the other hand, (Chowdhury et al.,2021)[12] present a review of 145 papers in the literature. We synthesize this work summarized in Figure 2, which presents the classification in question of a number of papers that study each problem's effects on SC during COVID-19. We conclude that the risk most addressed by researchers during the COVID period is the risk of demand.

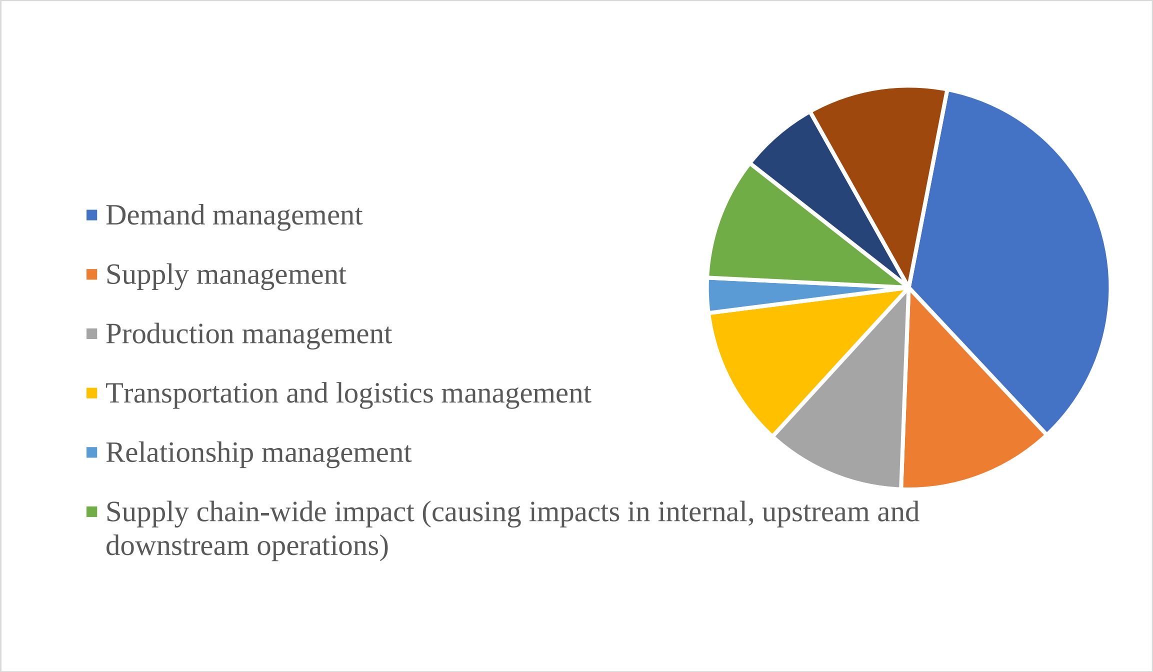


FIG. 2 – Types of problems found in the study of (Chowdhury et al.,2021)[12]

4 The role of mathematical modeling on SCM during COVID-19

In this section, our objective is to determine the role of mathematics modeling for SCM during COVID-19, In this study we search to find the response to this question :

Q: How can mathematical modeling create a resilient or robust supply chain during COVID-19?

Firstly, based on (Ozdemir et al., 2022) [29] definition of supply chain resilience: "Supply chain resilience indicates an ability to recover from an un desired performance level to a planned performance level by taking actions towards recovery or adaptation. Preparedness, alertness and agility are three pillars of supply chain resilience. They aim to minimize the effects of the disruption and ensure recovery as quickly as possible."

To obtain this objective we will follow the methodology illustrated in Figure 3.

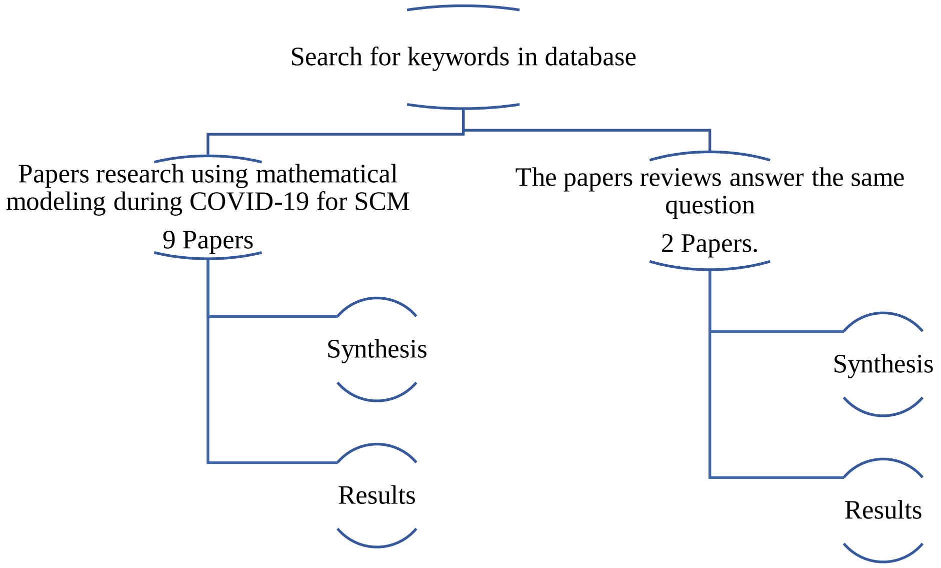


FIG. 3 – Study methodology and process

4.1 The paper's reviews answer the same question

Our question proposed in this study has become important for many researchers during COVID-19, hundreds of papers exist in the literature take it is a problem to search, after a look into the Google Scholar engine using the following keywords: SCM, COVID-19 and mathematical modeling. We found very satisfactory results and we chose the following items: (Chen et al.,2023)[11] and (Chen et al.,2023)[10].

The paper of the authors [11] proposed the following research questions:

- RQ1: What characteristics of CSc have been modeled quantitatively?
- RQ2: What are the various modeling approaches adopted when mathematical modeling and simulation are used in CSCM?
- RQ3: How frequently are different mathematical modeling, simulation and hybrid frameworks applied to model CSC integration and performance problems?

In this study, the question that converges towards our research question is the 3rd: it seeks to find the role of modeling in SCM of risks. Chen first followed the methodology of synthesizing 179 papers found in the database under the theme of construction supply chain. In the discussions and results section, the authors claim that most papers seeking to create chains of resilience during COVID-19 used mathematical modeling to achieve the objectives.

“Moreover, our analysis also suggests that production planning and project scheduling are two areas that can be researched in conjunction with CSC (construction supply chain) Integrations to enhance the study of the CSC. CSC operational risk management is a primary research area in improving project’s delivery process, and problems in this area are mainly addressed using mathematical modeling” (Chen et al.,2023)[11]

For the second paper, it is based on one of the components which is the same direction of research of our question and the following:

“Distribution Papers Based on Supply Chain Network Disruption during the COVID-19 Pandemic”.

The authors of this paper gathered 121 papers which are published during COVID-19, after a global synthesis they arrived at a result that confirms the role of mathematical modeling for SCM during COVID-19, he concludes this result:

"The SC disruption problem during the COVID-19 pandemic has also been studied in the literature, and various optimization models have been constructed to enhance SC resilience and mitigate SC disruption." (Chen et al.,2023) [10]

4.2 Papers research using mathematical modeling during COVID-19 for SCM

We have cited 9 research papers that use quantitative mathematical modeling methods to manage the supply chain during COVID-19 and mentioned with their results as follows:

To achieve a resilience chain the authors (Liu et al., 2023) [24], propose a reactive way for their management, using modeling to minimize the risk disruption of COVID-19, choosing the least expensive action to react with the risk and despite a new supply chain ripple effect management problem. The results of the solution of mathematical modeling given in this paper can create SC resilience during COVID-19, through the management of ripple effect.

A novel bi-objective Mixed-Integer Linear Programming (MILP) model is suggested in the paper by (Tirkolaee et al., 2023) [2] to formulate the problem which aims to minimize network costs and maximize job opportunities in SC of blood while considering the adverse effects of the pandemic COVID-19. The model resolution proposed by the authors can manage to minimize the load costs and verify work opportunities which means that their model creates a resilience chain during COVID-19.

The work of (Tang et al., 2022) [34] investigates a new multi-period vaccination planning problem that optimizes vaccination recipient's total travel distance (service level) and operational cost. An optimal plan determines, for each period, which vaccination sites to open, how many vaccination stations to launch at each site, how to assign recipients from different locations to opened sites, and the replenishment quantity of each site. Results from a case study indicate that their methods reduce the operational cost and the total travel distance by up to 9.3% and 36.6%, respectively. Managerial implications suggest enlarging the service capacity of vaccination sites can help improve the performance of the vaccination program.

(Azani et al., 2022) [5] their paper discusses a MINLP optimization model for handling the impact of the COVID-19 pandemic based on the food supply network through Food Hubs (FHs). In this research, the concept of FH has been used for a more effective and faster connection of consumers to production sites. Due to the prevention of the spread of coronavirus and the quarantine conditions, the areas have been divided into two parts (high-risk and low-risk) and two scenarios have been defined for this supply chain. (Azani et al., 2022) [5] confirm after solving the model and obtaining results. In this study, a practical solution has been presented to reduce the spread of the virus at the macro level and limit the communication of the regions of high and low risk.

In the study of (Paul et al., 2022) [12], we examine SC recovery for high-demand items (e.g., hand sanitizer and face masks). First, their developed a stochastic mathematical model to optimize recovery for a three-stage SC exposed to the multi-dimensional impacts of the COVID-19 pandemic. This allows us to generalize a novel problem setting with simultaneous demand, supply, and capacity uncertainty in a multi-stage SC recovery context. The findings in the results revealed that companies could significantly improve their total profit by implementing the strategies suggested for various scenarios and the model developed in this study.

Authors (Lozano-Diez et al., 2020) [25] develop a supply chain design model for a resilient supply network. The model is a mixed-integer linear problem with the objective function minimize cost and assumption of 2 scenarios, scenarios I and II are those that cause the greatest negative impact on the operation of the supply chain. In discussion and results finding the model mathematical in scenario 2 after the resolution, the highest profits are generated with a service level above 95%.

Healthcare supply chain is one of the chains that need to be resilient during COVID-19 because of the necessary need of customers for medical products especially in this period, the authors (Ala et al., 2024)[2] in their work propose a multi-objective mathematical model with different resolution approaches. The numerical results presented in this research confirm that the introduced fuzzy multi-objective model and its solution method can provide an advanced adaptation to the uncertain and complex conditions of resilient SC during the period of COVID-19.

(Shayannia et al., 2023)[32] development of multi-echelon and multi-product mathematical modeling to evaluate the pharmaceutical supply chain during the COVID-19 pandemic. The results of mathematical modeling used in this study can satisfy all the demands of the customers during COVID-19.

Blood is a product whose its demand is always increasing with delivery conditions in a very short period and the obligation to satisfy all demands, Fatemi's [15] study seeks to create a resilient blood supply chain during COVID-19. The mathematical model proposed by Fatemi is multi-objective with the objective of minimizing lost customer demands, this result shows that the management demonstrated in this study has won the challenge of building a resilient chain during COVID-19 to satisfy blood demands.

TAB. 4 – Solution methodology used by the papers cited in our study

Papers	Resolution Approach
(Liu et al., 2023) [24]	ϵ-constraint method non-dominated sorting genetic algorithm II
(Tirkolaee et al., 2023) [36]	Interactive possibilistic programming
(Tang et al., 2022) [34]	The weighted-sum approach ϵ-constraint method Non-Dominated Sorting Genetic Algorithm II (NSGA-II)
(Azani et al., 2022) [5]	ϵ-constraint method Genetic Optimization Algorithm
(Paul et al., 2022) [30]	Enhanced multi-operator differential evolution.
(Lozano-Diez et al., 2020) [25]	The anylogistix software uses CPLEX
(Ala et al., 2024)[2]	ϵ-constraint method Multi-objective gray wolf optimizer Nondominated sorting genetic algorithm II Multi-objective differential evolution algorithm Multi-objective methods evaluation metrics
(Shayannia et al., 2023)[32]	Multi-objective Particle Swarm Optimization Nondominated sorting genetic algorithm II
(Fatemi et al., 2022)[15]	LP_metric approach Goal attainment approach

Finally, all these papers used quantitative method mathematics modeling for multi-criteria decision-making, can achieve the objectives outlined, through the proposal of mathematical models with the objectives of reducing all risks or avoiding them during COVID-19, those risks that blocked one of the chain operations.

5 Conclusion and future directions

Although it is impossible to predict the onset of global crises such as the COVID-19 pandemic, companies can prepare to mitigate their impacts by developing robust and resilient supply chain (SC) processes and emergency preparedness plans. The effectiveness of these plans relies heavily on a thorough understanding of the pandemic's impact on Sc [17]. This paper synthesizes mathematical research to model the challenges caused by COVID-19 that can disrupt or block Sc, along with a review of two papers addressing similar research questions.

Mathematical models developed during COVID-19 have successfully strengthened SC resilience and helped companies and customers achieve their goals, ensuring the continuity of operations even amid the pandemic. This study highlights various types of mathematical models from the literature that address different disruptions in Sc, including those in the agri-food, pharmaceutical, and vaccination sectors. These models focus on managing demand, supply, production, transportation and logistics, relationships, SC-wide impacts (affecting internal, upstream, and downstream operations), financial management, and sustainability.

The pandemic underscored the critical role of mathematical models in achieving business and government objectives. This study provides a review of articles using mathematical modeling to manage supply chains during COVID-19 so that specialists can take advantage of it to avoid futuristic risks. Future research should build upon these models by incorporating emerging technologies like artificial intelligence and big data to strengthen SC defenses against future pandemics further.

References

[1] Abbasi, S., Daneshmand-Mehr, M., and Ghane Kanafi, A. Green Closed-Loop Supply Chain Network Design During the Coronavirus (COVID-19) Pandemic: A Case Study in the Iranian Automotive Industry, *Environ. Model*. Assess., 28(1): 69-103, 2023, doi: 10.1007/s10666-022-09863-0.

[2] Ala, A., Goli, A., Mirjalili, S., and Simic, V. A fuzzy multi-objective optimization model for sustainable healthcare supply chain network design, *Appl. Soft Comput*, 150: 111-012, 2024, doi: 10.1016/j.asoc.2023.111012.

[3] Alizadeh-Meghrazi, M., Tosarkani, B. M., Amin, S, H., Popovic, M, R., and Ahi, P., Design and optimization of a sustainable and resilient mask supply chain during the COVID-19 pandemic: A multi-objective approach, *Environ. Dev. Sustain*., 2022, doi: 10.1007/s10668-022-02604-z.

[4] Aqlan, F., and Lam, S. S. Supply chain risk modelling and mitigation, *Int. J. Prod. Res*., 53(18): 5640-5656, 2015, doi: 10.1080/00207543.2015.1047975.

[5] Azani, M., Shaerpour, M., Yazdani, M. A., Aghsami, A., and Jolai, F. A Novel Scenario- Based Bi-objective Optimization Model for Sustainable Food Supply Chain During the COVID-19: a Case Study, Process Integr. *Optim. Sustain*., 6(1):139-159, 2022, doi: 10.1007/s41660-021-00203-5.

[6] Bouti, N., Boukallal, I., and El Khoukhi, F. *Optimization of Inventory Management: A Literature Review*, Chapter and Conference Paper, Digital Technologies and Applications Proceedings of ICDTA’23, Part of the Lecture Notes in Networks and Systems book series (LNNS, volume 669), 2:923-933, 2023, doi: 10.1007/978-3-031-29860-8_92.

[7] Bouti, N., Boukallal, I., and El Khoukhi, F. *Sustainable Urban Logistics and Supply Chain 4.0: Literature Overview and Future Directions*, Chapter and Conference Paper, Proceeding of GOL'24. Smart Sustainable and Green Logistics, Part of the Lecture Notes in Systems book series (LNNS 1104), 1: 335-344, 2024, doi: 10.1007/978-3-031-68628-3_32.

[8] Bouti, N., and El Khoukhi, F. *Smart Supply Chain Management: A Literature Review*. Chapter and Conference Paper, Digital Technologies and Applications Proceedings of ICDTA’23, Part of the Lecture Notes in Networks and Systems book series (LNNS, vol. 668), 1:896-904, 2023, doi: 10.1007/978-3-031-29857-8_89.

[9] Boukallal, I., Bouti, N., and El Khoukhi, F. *Le problème AF-VRP : une revue de la littérature*, 5ème édition de la Conférence Internationale sur le Monitoring des Systèmes Industriels (CIMSI'22), 14-15, Décembre 2022, Tanger, Maroc.

[10] Chen, L., Dong, T., Peng, J., and Ralescu, D., Uncertainty Analysis and Optimization Modeling with Application to Supply Chain Management: A Systematic Review, *Mathematics*, 11(11): 2530, 2023, doi: 10.3390/math11112530.

[11] Chen, Z., and Hammad, A. W. A. Mathematical modelling and simulation in construction supply chain management, *Autom. Constr.*, 156: 105-147, 2023, doi: 10.1016/j.autcon.2023.105147.

[12] Chowdhury, P., Paul, S. K., Kaisar, S., and Moktadir, Md. A. COVID-19 pandemic related supply chain studies: A systematic review, Transp. Res. *Part E Logist. Transp*. Rev., 148 : 102-271, 2021, doi: 10.1016/j.tre.2021.102271.

[13] El Baz, J., and Ruel, S. Can supply chain risk management practices mitigate the disruption impacts on supply chains' resilience and robustness? Evidence from an empirical survey in a COVID-19 outbreak era, *Int. J. Prod. Econ*.,233:107-972, 2021, doi: 10.1016/j.ijpe.2020.107972.

[14] El Khayati, M., Ouba, A., Elkettani, Y., and El Khoukhi, F. *A new ground truth dataset for the automatic validation of Arabic handwritten character segmentation algorithms*, 6th International Conference on Networking, Intelligent Systems and Security (NISS'23), May 24-25, 2023, Larache, Maroc, 2023.

[15] Fatemi, M. S., Ghodratnama, A., Tavakkoli-Moghaddam, R., and Kaboli, A. A multi-functional tri-objective mathematical model for the pharmaceutical supply chain considering congestion of drugs in factories, Res. *Transp. Econ.*, 92:101-094, 2022, doi: 10.1016/j.retrec.2021.101094.

[16] Goodarzian, F., Kumar, V., and Ghasemi, P. Investigating a citrus fruit supply chain network considering CO2 emissions using meta-heuristic algorithms, *Ann. Oper*. Res., 2022, doi: 10.1007/s10479-022-05005-7.

[17] Hamdi, F., Ghorbel, A., Masmoudi, F., and Dupont, L. Optimization of a supply portfolio in the context of supply chain risk management: literature review, *J. Intell. Manuf*., 29(4):763-788, 2018, doi: 10.1007/s10845-015-1128-3.

[18] Hosseini, S., Ivanov, D., and Dolgui, A. Review of quantitative methods for supply chain resilience analysis, Transp. Res. *Part E Logist. Transp*. Rev.,125 :285-307, 2019, doi: 10.1016/j.tre.2019.03.001.

[19] Ho, W., Zheng, T., Yildiz, H., and Talluri, S. Supply chain risk management: a literature review, *Int. J. Prod*. Res., 53(16):5031-5069, 2015, doi: 10.1080/00207543.2015.1030467.

[20] Ivanov, D., and Dolgui, A. Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak, Int. J. Prod. Res., 58(10):2904-2915, 2020, doi: 10.1080/00207543.2020.1750727.

[21] Ivanov, D., Dolgui, A., Blackhurst, J. V., and Choi, T.-M. Toward supply chain viability theory: from lessons learned through COVID-19 pandemic to viable ecosystems, *Int. J. Prod*. Res., 61(8):2402-2415, 2023, doi: 10.1080/00207543.2023.2177049.

[22] Ivanov, D., Chen, W., Coit, D. W., and Altay, N. Modeling and Optimization of Supply Chain Resilience to Pandemics and Long-Term Crises, *IISE Trans.*, 56(7):683-684, 2024, doi: 10.1080/24725854.2024.2317845.

[23] Kleindorfer, P. R., and Saad, G. H. Managing Disruption Risks in Supply Chains, Prod. *Oper. Manag*., 14(1): 53-68, 2009, doi: 10.1111/j.1937-5956. 2005.tb00009. x.

[24] Liu, M., Lin, T., Chu, F., Ding, Y., Zheng, F., and Chu, C. Bi-objective optimization for supply chain ripple effect management under disruption risks with supplier actions, *Int. J. Prod. Econ.*, 265:108-997, 2023, doi: 10.1016/j.ijpe.2023.108997.

[25] Lozano-Diez, J., Marmolejo-Saucedo, J., and Rodriguez-Aguilar, R. Designing a resilient supply chain: An approach to reduce drug shortages in epidemic outbreaks, *EAI Endorsed Trans. Pervasive Health Technol.*,6(21):164-260, 2020, doi: 10.4108/eai.13-7-2018.164260.

[26] Montoya-Torres, J. R., Muñoz-Villamizar, A., and Mejia-Argueta, C. Mapping research in logistics and supply chain management during COVID-19 pandemic, *Int. J. Logist. Res. Appl.*, 26(4):421-441, 2023, doi: 10.1080/13675567.2021.1958768.

[27] Nooraie, V., Fathi, M., Narenji, M., Parast, M. M., Pardalos, P. M., and Stanfield, P. M. A multi-objective model for risk mitigating in supply chain design, *Int. J. Prod. Res.*, 58(5):1338-1361, 2020, doi: 10.1080/00207543.2019.1633024.

[28] Ouba, A., and El Khoukhi, F. *Supply chain risk management: A literature review*, International Conference on Services and Industry of the Future (ICSIF2023), October, 24-26, Rabat, Morocco,2023.

[29] Ozdemir, D., Sharma, M., Dhir, A., and Daim, T. Supply chain resilience during the COVID-19 pandemic, *Technol. Soc.*, 68:101-847, 2022, doi: 10.1016/j.techsoc.2021.101847.

[30] Paul, S. K., Chowdhury, P., Chakrabortty, R. K., Ivanov, D., and Sallam, K. (2022), A mathematical model for managing the multi-dimensional impacts of the COVID-19 pandemic in supply chain of a high-demand item, *Ann. Oper*. Res., 2022, doi: 10.1007/s10479-022-04650-2.

[31] Schleifenheimer, M., and Ivanov, D. Pharmaceutical retail supply chain responses to the COVID-19 pandemic, *Ann. Oper*. Res.,2024, doi: 10.1007/s10479-024-05866-0.

[32] Shayannia, S, A., Presenting an agile supply chain mathematical model for COVID-19 (Corona) drugs using metaheuristic algorithms (case study: pharmaceutical industry) , *Environ. Sci. Pollut. Res.*, 30(3): 6559-6572, 2023, doi: 10.1007/s11356-022-22608-6.

[33] Suryawanshi, P., and Dutta, P. Optimization models for supply chains under risk, uncertainty, and resilience: A state-of-the-art review and future research directions, Transp. Res. *Part E Logist. Transp*. Rev., 157 :102 553, 2022, doi: 10.1016/j.tre.2021.102553.

[34] Tang, L., Li, Y., Bai, D., Liu, T., and Coelho, L. C. Bi-objective optimization for a multi-period COVID-19 vaccination planning problem, *Omega*, 110:102-617, 2022, doi: 10.1016/j.omega.2022.102617.

[35] Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., and Fischl, M. Artificial intelligence in supply chain management: A systematic literature review, *J. Bus. Res.*, 122:502-517, 2021, doi: 10.1016/j.jbusres.2020.09.009.

[36] Tirkolaee, E. B., Golpîra, H., Javanmardan, A., and Maihami, R. A socio-economic optimization model for blood supply chain network design during the COVID-19 pandemic: An interactive possibilistic programming approach for a real case study, *Socioecon. Plann.* Sci., 85:101-439, 2023, doi: 10.1016/j.seps.2022.101439.

[37] Vafadarnikjoo, A., Moktadir, Md. A., Paul, S. K., and Ali, S. M. A novel grey multi-objective binary linear programming model for risk assessment in supply chain management, *Supply Chain Anal.*, 2:100-012, 2023, doi: 10.1016/j.sca.2023.100012.

[38] Wang, X., Chen, G., and Xu, S. Bi-objective green supply chain network design under disruption risk through an extended NSGA-II algorithm, *Clean. Logist. Supply Chain*, 3: 100-025, 2022, doi: 10.1016/j.clscn.2021.100025.

[39] Xu, Z., Elomri, A., Kerbache, L., and El Omri, A. *Impacts of COVID-19 on Global Supply Chains: Facts and Perspectives*, IEEE Eng. Manag. Rev., 48(3): 153-166, September 2020, doi: 10.1109/EMR.2020.3018420.